

# Optimizing Performance of Continuous-Time Stochastic Systems using Timeout Synthesis<sup>★</sup>

Tomáš Brázdil<sup>1</sup>, L'uboš Korenčiak<sup>1</sup>, Jan Krčál<sup>2</sup>, Petr Novotný<sup>3</sup>, and Vojtěch Řehák<sup>1</sup>

<sup>1</sup> Faculty of Informatics, Masaryk University, Brno, Czech Republic  
{brazdil, korenciak, rehak}@fi.muni.cz

<sup>2</sup> Saarland University – Computer Science, Saarbrücken, Germany  
krca1@cs.uni-saarland.de

<sup>3</sup> IST Austria, Klosterneuburg, Austria  
petr.novotny@ist.ac.at

**Abstract.** We consider parametric version of fixed-delay continuous-time Markov chains (or equivalently deterministic and stochastic Petri nets, DSPN) where fixed-delay transitions are specified by parameters, rather than concrete values. Our goal is to synthesize values of these parameters that, for a given cost function, minimize expected total cost incurred before reaching a given set of target states. We show that under mild assumptions, optimal values of parameters can be effectively approximated using translation to a Markov decision process (MDP) whose actions correspond to discretized values of these parameters. To this end we identify and overcome several interesting phenomena arising in systems with fixed delays.

## 1 Introduction

Continuous-time Markov chains (CTMC) are a fundamental model of stochastic systems with discrete state-spaces that evolve in continuous-time. Several higher level modelling formalisms, such as stochastic Petri nets and stochastic process algebras, use CTMC as their semantics. As such, CTMC have been applied in performance and dependability analysis in various contexts ranging from aircraft communication protocols (see, e.g. [36]) to models of biochemical systems (see, e.g. [23]).

There are several equivalent definitions of CTMC (see, e.g. [15,32]). We may define a (uniformized, finite-state) CTMC to consist of a finite set of states  $S$  coupled with a common rate  $\lambda$  and a stochastic matrix  $P \in \mathbb{R}_{\geq 0}^{S \times S}$  specifying probabilities of transitions between states. An execution starts in a given initial state. In every step, the CTMC waits for a duration that is selected randomly according to the exponential distribution with the rate  $\lambda$ , and then moves to a state  $s'$  randomly chosen with probability  $P(s, s')$ .

The practical interpretation of the above semantics is that in every state the system waits for an event to occur and then reacts by changing its state. A typical example is a model of

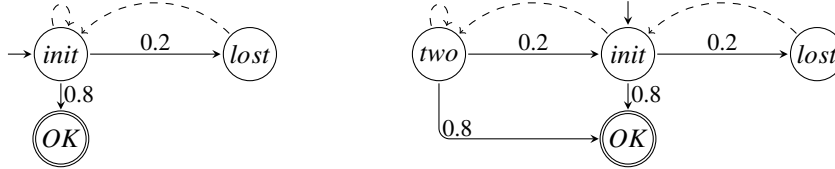
<sup>★</sup> The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement n° [291734]. This work is partly supported by the German Research Council (DFG) as part of the Transregional Collaborative Research Center AVACS (SFB/TR 14), by the EU 7th Framework Programme under grant agreement no. 295261 (MEALS) and 318490 (SENSATION), by the Czech Science Foundation, grant No. 15-17564S, and by the CAS/SAFEA International Partnership Program for Creative Research Teams.

a simple queue to which new customers come in random intervals and are also served in random intervals. However, in practice, events are usually not exponentially distributed, and, in fact, their distributions may be quite far from being exponential. To deal with such events, phase-type approximation technique [31] is usually applied. Unfortunately, as already noted in [31], some distributions cannot be efficiently fit with phase-type approximation. In particular, degenerate distributions of events with fixed delays, i.e., events that occur after a fixed amount of time with probability 1, form a distinguished example of this phenomenon (for more details see [26]). However, as events with fixed delays play a crucial role in many systems, especially in communication protocols [33], time-driven real-time scheduling [35], etc., they should be handled faithfully in modelling and analysis.

Inspired by deterministic and stochastic Petri nets [29] and delayed CTMC [16] with at most one non-exponential transition enabled in any time, we study fixed-delay CTMC (fdCTMC), the CTMC extended with *fixed-delay transitions*. More concretely, we specify a set of states  $S_{fd} \subseteq S$  where fixed-delay transitions are enabled and add a stochastic matrix  $F \in \mathbb{R}_{\geq 0}^{S_{fd} \times S}$  specifying probabilities of fixed-delay transitions between states. In addition, we consider a *delay function*  $\mathbf{d} : S_{fd} \rightarrow \mathbb{R}_{>0}$ . The semantics can be intuitively described as follows. Imagine a CTMC extended with an alarm clock. At the beginning of an execution, the alarm clock is turned off and the process behaves as the original CTMC. Whenever a state  $s$  of  $S_{fd}$  is visited and the alarm clock is off at the time, it is turned on and set to ring after  $\mathbf{d}(s)$  time units. Subsequently, the process keeps behaving as the original CTMC until either a state of  $S \setminus S_{fd}$  is visited (in which case the alarm clock is turned off), or the alarm clock rings in a state  $s'$  of  $S_{fd}$ . In the latter case, a fixed-delay transition takes place, which means that the process changes the state randomly according to the distribution  $F(s', \cdot)$ , and the alarm clock is either turned off or newly set (when entering a state of  $S_{fd}$ ).

In most practical applications mentioned above, fixed-delay transitions are determined by the design of the system and often strongly influence performance of the system. Indeed, both timeouts in network protocols as well as scheduling intervals in real-time systems directly influence performance of the respective systems and their manual setting usually requires considerable effort and expertise. This motivates us to consider the fixed-time delays  $\mathbf{d}(s)$  as *free parameters* of the model, and develop techniques for their optimization with respect to a given performance measure.

*Example 1.* We demonstrate the concept on two different models of sending *one* segment of data in the *alternating bit protocol*. In the protocol, each segment of data is retransmitted until an acknowledgement is received. The delay between retransmissions has impact on throughput of the protocol as well as on network congestion. In the simpler model below on the left, the data is sent in state *init*. The exp-delay transitions, drawn as solid arrows, model message loss (with probability 0.2) and delivery (with probability 0.8). For simplicity we use rate 1 and omit self loops of exponential transitions in all examples. The fixed-delay transitions, drawn as dashed arrows, cause the data to be retransmitted. Note that whenever the data is retransmitted, the previous message with the data is canceled in this model.



The more faithful model on the right models up to two messages with the data segment being delivered concurrently. For choosing an optimal delay between retransmissions, we need to formalize how to express performance of the protocol.

To express performance properties, we use standard cost (or reward) structures (see, e.g. [34]) that assign numerical rewards to states and transitions. More precisely, we consider the following three cost functions:  $\mathcal{R} : S \rightarrow \mathbb{R}_{\geq 0}$ , which assigns a cost rate  $\mathcal{R}(s)$  to every state  $s$  so that the cost  $\mathcal{R}(s)$  is paid for every unit of time spent in the state  $s$ , and functions  $\mathcal{I}_P, \mathcal{I}_F : S \times S \rightarrow \mathbb{R}_{\geq 0}$  that assign to each exp-delay and fixed-delay transition, respectively, the cost that is immediately paid when the transition is taken. Note that  $\mathcal{R}$  is usually used to express time spent in individual states, while the other two cost functions are used to quantify the difficulty of dealing with events corresponding to transitions. The performance measure itself is the *expected total cost incurred before reaching a given set of states  $G$  starting in a given initial state  $s_{in}$* . For this moment, let us denote this measure by  $E_{\mathbf{d}}$ , stressing the fact that it depends on the delay function  $\mathbf{d}$  which is the only variable quantity in our optimization task:

*Problem 1 (Cost optimization).* For a subspace of delay functions  $D \subseteq (\mathbb{R}_{>0})^{S_{fd}}$  and a given approximation error  $\varepsilon > 0$ , compute a delay function  $\mathbf{d} \in D$  that is  $\varepsilon$ -optimal within  $D$ , i.e.

$$\left| \inf_{\mathbf{d}' \in D} E_{\mathbf{d}'} - E_{\mathbf{d}} \right| < \varepsilon.$$

*Example 1 (cont.)* We can model the expected cost of sending one data segment in our examples as follows. To take into account the expected time of data delivery, we set the cost rate of each state to, e.g., 1. To take into account the expected number of retransmissions, we set the cost of each fixed-delay transition, e.g., to 3. The cost of each exp-delay transition is set to 0. Now the goal for the left model is to find a delay  $\mathbf{d}(init)$  optimizing the expected total cost incurred before reaching the state *OK*. Note that  $\mathbf{d}$  is never set in the state *lost*. The goal is the same for the model on the right where  $\mathbf{d}$  is set also in the state *two*. Note that it makes no sense to synthesize different delays  $\mathbf{d}(init)$  and  $\mathbf{d}(two)$  as the states *init* and *two* are indistinguishable in the implementation of the protocol. Therefore, we need to require that the synthesised delay function satisfies  $\mathbf{d}(init) = \mathbf{d}(two)$ .

*Our contribution:* We consider fixed-delay CTMC as a natural extension of CTMC suitable for algorithmic synthesis of fixed timeouts. Upon this model, we investigate algorithmic complexity of the cost optimization problem. This is, to the best of our knowledge, the most general attempt at fully automatic synthesis of timeouts in continuous-time stochastic systems. We provide algorithms for solving the following two special cases of the cost optimization problem under the assumption that the reward rate  $\mathcal{R}(s)$  is *positive* in every state  $s$ :

1. **Unconstrained optimization** where we demand  $D = (\mathbb{R}_{>0})^{S_{fd}}$ , i.e. the set of all delay functions. We solve this problem by reduction to a finite Markov decision process (MDP) whose actions correspond to *discretized* (i.e. rounded onto a finite mesh) values of delays in the individual states, and then apply standard polynomial time algorithms for synthesis of the delays (note that a brute force search through a "discretized" subset of  $D$  would be exponentially worse). The most non-trivial part is to prove

that the delays may be discretized. We show that a naïve rounding of a near-optimal delay function may cause arbitrarily high *absolute* error. Our solution, based on rather non-trivial insights into the structure of fdCTMCs, avoids this obstacle by identifying "safe" delay functions that may be rounded with an error bounded (exponentially) in the size of the system. This leads to an exponential time algorithm for solving the cost optimization problem.

2. **Bounded optimization under partial observation** where we introduce bounds  $\underline{d}, \bar{d} > 0$  together with an equivalence relation  $\equiv$  on  $S_{fd}$  and demand  $D$  to be the set of all delay functions  $\mathbf{d}$  satisfying the following conditions:

- $\underline{d} \leq \mathbf{d}(s) \leq \bar{d}$  for all  $s \in S_{fd}$ ,
- $\mathbf{d}(s) = \mathbf{d}(s')$  whenever  $s \equiv s'$ .

Like in the Example 1, the equivalence  $\equiv$  can be used to hide information about detailed internal structure of states which is often needed in practical applications. In this paper, we show that the bounded optimization under partial observation can be solved in time doubly exponential in  $\bar{d}$  and exponential in all other parameters.

We also consider the corresponding approximate threshold variant: For a given  $x$  decide whether  $\inf_{\mathbf{d} \in D} E_{\mathbf{d}} > x + \varepsilon$ , or  $\inf_{\mathbf{d} \in D} E_{\mathbf{d}} < x - \varepsilon$  (for  $\inf_{\mathbf{d} \in D} E_{\mathbf{d}} \in [x - \varepsilon, x + \varepsilon]$  an arbitrary answer may be given). We show that this bounded optimization problem is NP-hard, thus a polynomial time solution of the bounded optimization under partial observation is unlikely.

The assumption that all delays are between fixed thresholds  $\underline{d}$  and  $\bar{d}$  is crucial in our approach. As we discuss in Section 4, without this assumption the optimization under partial observation becomes much trickier and we leave its solution for future work.

*Related work.* Various forms of continuous-time stochastic processes with fixed-delay transitions have already been studied, see e.g. [29,13,1,9,4]. In particular, as noted above, our definition of fdCTMC is closely related to the original definition of deterministic and stochastic Petri nets [29]. Papers on verification of continuous-time systems with timed automata (TA) specifications [12,4,5] are also related to our work as the constraints in timed automata resemble fixed-delay transitions. None of these works, however, considers synthesis of fixed-delays (or other parameters).

Parameter synthesis techniques have been developed for several models, such as parametric timed automata [2], parametric one-counter automata [17], parametric Markov models [18], etc. In continuous-time stochastic systems, [19,23] study synthesis of rates in CTMC which is a problem orthogonal to timeouts. Furthermore, optimal control of continuous-time (Semi)-Markov decision processes [30,10,8,6] can be viewed as synthesis of *discrete* parameters in continuous-time systems.

The problem of synthesizing timeouts as *continuous* parameters has been studied in variety of engineering contexts such as vehicle communication systems [25] and avionic subsystems [3,36]. To the best of our knowledge, no generic framework for synthesis of timeouts in stochastic continuous-time systems has been developed so far. In theoretical literature, only simpler cases have been addressed. For instance [11,37] consider a finite test case, a sequence of input and output actions, and synthesize times for input actions that maximize the probability of executing this acyclic sequence. Allowing cycles in fdCTMC makes the timeout synthesis problem much more demanding, e.g., due to potentially unbounded number of stochastic events between timeouts. Instead of static timeouts, [27,22] consider synthesis of "dynamic" timeouts where the delay is chosen based on the

history of the execution so far. Consequently, the delay can be changed *while it is elapsing* whenever stochastic events occur. This makes it much simpler to solve and also adequate for a different application domain.

Section 2 introduces fixed-delay CTMC and cost structures. Section 3 and Section 4 are devoted to unconstrained optimization and bounded optimization under partial observation, respectively. Due to space constraints, full proofs are in [7].

## 2 Preliminaries

We use  $\mathbb{N}_0$ ,  $\mathbb{R}_{\geq 0}$ , and  $\mathbb{R}_{> 0}$  to denote the set of all non-negative integers, non-negative real numbers, and positive real numbers, respectively. Furthermore, for a countable set  $A$ , we denote by  $\mathcal{D}(A)$  the set of discrete probability distributions over  $A$ , i.e. functions  $\mu : A \rightarrow \mathbb{R}_{\geq 0}$  such that  $\sum_{a \in A} \mu(a) = 1$ . Encoding size of an object  $O$  is denoted by  $\|O\|$ .

**Definition 1.** A fixed-delay CTMC structure (*fdCTMC structure*)  $C$  is a tuple  $(S, \lambda, P, S_{\text{fd}}, F, s_{\text{in}})$  where

- $S$  is a finite set of states,
- $\lambda \in \mathbb{R}_{> 0}$  is a (common) rate of exp-delay transitions,
- $P : S \times S \rightarrow \mathbb{R}_{\geq 0}$  is a stochastic matrix specifying probabilities of exp-delay transitions,
- $S_{\text{fd}} \subseteq S$  is a set of states where fixed-delay transitions are enabled,
- $F : S_{\text{fd}} \times S \rightarrow \mathbb{R}_{\geq 0}$  is a stochastic matrix specifying probabilities of fixed-delay transitions, and
- $s_{\text{in}} \in S$  is an initial state.

A fixed-delay CTMC (*fdCTMC*) is a pair  $C(\mathbf{d}) = (C, \mathbf{d})$  where  $C$  is a fdCTMC structure and  $\mathbf{d} : S_{\text{fd}} \rightarrow \mathbb{R}_{> 0}$  is a delay function which to every state where fixed-delay transitions are enabled assigns a positive delay.

A configuration of a fdCTMC is a pair  $(s, d)$  where  $s \in S$  is the current state and  $d \in \mathbb{R}_{> 0} \cup \{\infty\}$  is the remaining time before a fixed-delay transition takes place. We assume that  $d = \infty$  iff  $s \notin S_{\text{fd}}$ . To simplify notation, we similarly extend any delay function  $\mathbf{d}$  to all states  $S$  by assuming  $\mathbf{d}(s) = \infty$  iff  $s \notin S_{\text{fd}}$ .

An execution of  $C(\mathbf{d})$  starts in the configuration  $(s_0, d_0)$  with  $s_0 = s_{\text{in}}$  and  $d_0 = \mathbf{d}(s_{\text{in}})$ . In every step, assuming that the current configuration is  $(s_i, d_i)$ , the fdCTMC waits for some time  $t_i$  and then moves to a next configuration  $(s_{i+1}, d_{i+1})$  determined as follows:

- First, a waiting time  $t_{\text{exp}}$  for exp-delay transitions from  $s_i$  is chosen randomly according to the exponential distribution with the rate  $\lambda$ .
- Then
  - If  $t_{\text{exp}} < d_i$ , then an exp-delay transition occurs, which means that  $t_i = t_{\text{exp}}$ ,  $s_{i+1}$  is chosen randomly with probability  $P(s_i, s_{i+1})$ , and  $d_{i+1}$  is determined by

$$d_{i+1} = \begin{cases} d_i - t_{\text{exp}} & \text{if } s_{i+1} \in S_{\text{fd}} \text{ and } s_i \in S_{\text{fd}} \text{ (previous delay remains),} \\ \mathbf{d}(s_{i+1}) & \text{if } s_{i+1} \notin S_{\text{fd}} \text{ or } s_i \notin S_{\text{fd}} \text{ (delay is newly set or disabled).} \end{cases}$$

- If  $t_{\text{exp}} \geq d_i$ , then a fixed-delay transition occurs, which means that  $t_i = d_i$ ,  $s_{i+1}$  is chosen randomly with probability  $F(s_i, s_{i+1})$ , and  $d_{i+1} = \mathbf{d}(s_{i+1})$ .

This way, the execution of a fdCTMC forms a *run*, an alternating sequence of configurations and times  $(s_0, d_0)t_0(s_1, d_1)t_1 \dots$ . The probability measure  $\Pr_{C(\mathbf{d})}$  over runs of  $C(\mathbf{d})$  is formally defined in Appendix B.1.

*Total cost before reaching a goal* To allow formalization of performance properties, we enrich the model in a standard way (see, e.g. [34]) with costs (or rewards). A *cost structure* over a fdCTMC structure  $C$  with state space  $S$  is a tuple  $Cost = (G, \mathcal{R}, \mathcal{I}_P, \mathcal{I}_F)$  where  $G \subseteq S$  is a set of goal states,  $\mathcal{R} : S \rightarrow \mathbb{R}_{\geq 0}$  assigns a cost rate to every state, and  $\mathcal{I}_P, \mathcal{I}_F : S \times S \rightarrow \mathbb{R}_{\geq 0}$  assign an impulse cost to every exp-delay and fixed-delay transition, respectively. Slightly abusing the notation, we denote by  $Cost$  also the random variable assigning to each run  $\omega = (s_0, d_0)t_0 \dots$  the *total cost before reaching  $G$*  (in at least one transition), given by

$$Cost(\omega) = \begin{cases} \sum_{i=0}^{n-1} (t_i \cdot \mathcal{R}(s_i) + \mathcal{I}_i(\omega)) & \text{for minimal } n > 0 \text{ such that } s_n \in G, \\ \infty & \text{if there is no such } n, \end{cases}$$

where  $\mathcal{I}_i(\omega)$  equals  $\mathcal{I}_P(s_i, s_{i+1})$  for an exp-delay transition, i.e. when  $t_i < d_i$ , and equals  $\mathcal{I}_F(s_i, s_{i+1})$  for a fixed-delay transition, i.e. when  $t_i = d_i$ .

We denote the expectation of  $Cost$  with respect to  $\Pr_{C(\mathbf{d})}$  simply by  $E_{C(\mathbf{d})}$ , or by  $E_{C(\mathbf{d})}^{Cost}$  when  $Cost$  is not clear from context. Our aim is to (approximatively) minimize the expected cost, i.e. to find a delay function  $\mathbf{d}$  such that  $E_{C(\mathbf{d})} \leq Val[C] + \varepsilon$  where  $Val[C]$  denotes the optimal cost  $\inf_{\mathbf{d}'} E_{C(\mathbf{d}')}$ .

*Non-parametric analysis* Due to [28], we can easily analyze a fdCTMC where the delay function is *fixed*. Hence, both the expected total cost before reaching  $G$  and the reaching probabilities of states in  $G$  can be efficiently approximated.

**Proposition 1.** *There is a polynomial-time algorithm that for a given fdCTMC  $C(\mathbf{d})$ , cost structure  $Cost$  with goal states  $G$ , and an approximation error  $\varepsilon > 0$  computes  $x \in \mathbb{R}_{>0} \cup \{\infty\}$  and  $p_s \in \mathbb{R}_{>0}$ , for each  $s \in G$ , such that*

$$\left| E_{C(\mathbf{d})} - x \right| < \varepsilon \quad \text{and} \quad \left| \Pr_{C(\mathbf{d})}(\Diamond_G^s) - p_s \right| < \varepsilon$$

where  $\Diamond_G^s$  is the set of runs that reach  $s$  as the first state of  $G$  (after at least one transition).

*Markov decision processes.* In Section 3 we use a reduction of fdCTMC to discrete-time Markov decision processes (DTMDP, see e.g. [34]) with uncountable space of actions.

**Definition 2.** A DTMDP is a tuple  $\mathcal{M} = (V, Act, T, v_{in}, V')$ , where  $V$  is a finite set of vertices,  $Act$  is a (possibly uncountable) set of actions,  $T : V \times Act \rightarrow \mathcal{D}(V) \cup \{\perp\}$  is a transition function,  $v_{in} \in V$  is an initial vertex, and  $V' \subseteq V$  is a set of goal vertices.

An action  $a$  is *enabled* in a vertex  $v$  if  $T(v, a) \neq \perp$ . A *strategy* is a function  $\sigma : V \rightarrow Act$  which assigns to every vertex  $v$  an action enabled in  $v$ . The behaviour of  $\mathcal{M}$  with a fixed strategy  $\sigma$  can be intuitively described as follows: A run starts in the vertex  $v_{in}$ . In every step, assuming that the current vertex is  $v$ , the process moves to a new vertex  $v'$  with probability  $T(v, \sigma(v))(v')$ . Every strategy  $\sigma$  uniquely determines a probability measure  $\Pr_{\mathcal{M}(\sigma)}$  on the set of *runs*, i.e. infinite alternating sequences of vertices and actions  $v_0 a_1 v_1 a_2 v_2 \dots \in (V \cdot Act)^\omega$ ; see Appendix B.3 for details.

Analogously to fdCTMC, we can endow a DTMDP with a *cost function*  $A : V \times Act \rightarrow \mathbb{R}_{\geq 0}$ . We then define for each run  $v_0 a_1 v_1 a_2 \dots$ , the *total cost incurred before reaching  $V'$*  as  $\sum_{i=0}^{n-1} A(v_i, a_{i+1})$  if there is a minimal  $n > 0$  such that  $v_n \in V'$ , and as  $\infty$  otherwise. The

expectation of this cost w.r.t.  $\Pr_{\mathcal{M}(\sigma)}$  is similarly denoted by  $E_{\mathcal{M}(\sigma)}$  or by  $E_{\mathcal{M}(\sigma)}^A$  if the cost function is not clear from context.

Given  $\varepsilon \geq 0$ , we say that a strategy  $\sigma$  is  $\varepsilon$ -optimal in  $\mathcal{M}$  if  $E_{\mathcal{M}(\sigma)} \leq \text{Val}[\mathcal{M}] + \varepsilon$  where  $\text{Val}[\mathcal{M}] = \inf_{\sigma'} E_{\mathcal{M}(\sigma')}$ ; we call it *optimal* if it is 0-optimal. For any  $s \in S$ , let us denote by  $\mathcal{M}[s]$  the DTMDP obtained from  $\mathcal{M}$  by replacing the initial state by  $s$ . We call a strategy *globally ( $\varepsilon$ )-optimal* if it is ( $\varepsilon$ )-optimal in  $\mathcal{M}[s]$  for every  $s \in S$ . Sometimes, we restrict to a subset  $D$  of all strategies and denote by  $\text{Val}[\mathcal{M}, D]$  the restricted infimum  $\inf_{\sigma' \in D} E_{\mathcal{M}(\sigma')}$ .

### 3 Unconstrained Optimization

**Theorem 1.** *There is an algorithm that given a fdCTMC structure  $C$ , a cost structure  $\text{Cost}$  with  $\mathcal{R}(s) > 0$  for all  $s \in S$ , and  $\varepsilon > 0$  computes in exponential time a delay function  $\mathbf{d}$  with*

$$\left| E_{C(\mathbf{d})} - \inf_{\mathbf{d}'} E_{C(\mathbf{d}')} \right| < \varepsilon.$$

The rest of this section is devoted to a proof of Theorem 1, which consists of two parts. First, we reduce the optimization problem in the fdCTMC to an optimization problem in a *discrete time* Markov decision process (DTMDP) with uncountably many actions. Second, we present the actual approximation algorithm based on a straightforward discretization of the space of actions of the DTMDP. However, the proof of its error bound is actually quite intricate. The time complexity is exponential because the discretized DTMDP needs exponential size in the worst case. For complexity with respect to various parameters see Appendix C.5.

For the rest of this section we fix a fdCTMC structure  $C = (S, \lambda, P, S_{\text{fd}}, F, s_{\text{in}})$ , a cost structure  $\text{Cost} = (G, \mathcal{R}, \mathcal{I}_P, \mathcal{I}_F)$ , and  $\varepsilon > 0$ . We assume that  $\text{Val}[C] < \infty$ . The opposite case can be easily detected by fixing an arbitrary delay function  $\mathbf{d}$  and finding out whether  $E_{C(\mathbf{d})} = \infty$  by Proposition 1. This is equivalent to  $\text{Val}[C] = \infty$  by the following observation.

**Lemma 1.** *For any delay functions  $\mathbf{d}, \mathbf{d}'$  we have  $E_{C(\mathbf{d})} = \infty$  if and only if  $E_{C(\mathbf{d}')} = \infty$ .*

To further simplify our presentation, we assume that each state  $s$  of  $S_{\text{fd}}$  directly encodes whether the delay needs to be reset upon entering  $s$ . Formally, we assume  $S_{\text{fd}} = S^{\text{reset}} \uplus S^{\text{keep}}$  where  $s' \in S^{\text{reset}}$  if  $P(s, s') > 0$  for some  $s \in S \setminus S_{\text{fd}}$  or if  $F(s, s') > 0$  for some  $s \in S_{\text{fd}}$ ; and  $s' \in S^{\text{keep}}$  if  $P(s, s') > 0$  for some  $s \in S_{\text{fd}}$ . We furthermore assume that  $s_{\text{in}} \in S^{\text{reset}}$  if  $s_{\text{in}} \in S_{\text{fd}}$ . Note that each fdCTMC structure can be easily transformed to satisfy this assumption in polynomial time by duplication of states  $S_{\text{fd}}$ , see, e.g., Example 2.

#### 3.1 Reduction to DTMDP $\mathcal{M}$ with Uncountable Space of Actions.

We reduce the problem into a *discrete-time* problem by capturing the evolution of the fdCTMC only at *discrete moments* when a transition is taken after which the fixed-delay is (a) newly set, or (b) switched off, or (c) irrelevant as the goal set is reached. This happens exactly when one of the states of  $S' = S^{\text{reset}} \cup (S \setminus S_{\text{fd}}) \cup G$  is reached. We define a DTMDP  $\mathcal{M} = (S', \text{Act}, T, s_{\text{in}}, G)$  with a cost function  $\mathbb{E}$ :

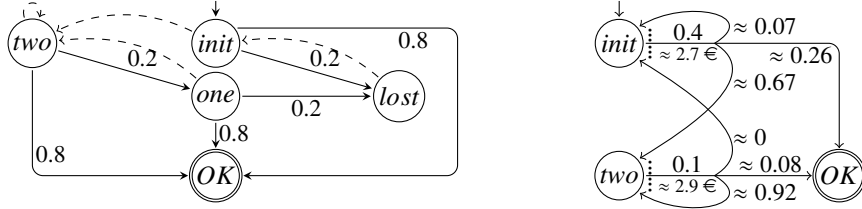
- $\text{Act} := \mathbb{R}_{>0} \cup \{\infty\}$ ; where actions  $\mathbb{R}_{>0}$  are enabled in  $s \in S^{\text{reset}}$  and action  $\infty$  is enabled in  $s \in S \setminus S_{\text{fd}}$ .

- Let  $s \in S'$  and  $d$  be an action of  $\mathcal{M}$ . Intuitively, we define  $T(s, d)$  and  $\mathbb{E}(s, d)$  to summarize the behaviour of the fdCTMC starting in the configuration  $(s, d)$  until the first moment when  $S'$  is reached again. Formally, let  $C[s](d)$  denote a fdCTMC obtained from  $C$  by changing initial state to  $s$  and fixing a delay function that assigns  $d$  to  $s$  (and arbitrary values elsewhere). We define  $\mathbb{E}(s, d)$  as the cost accumulated before reaching another state of  $S'$  and  $T(s, d)(s')$  as the probability that  $s'$  is the first such a reached state of  $S'$ . That is,

$$\mathbb{E}(s, d) = E_{C[s](d)}^{Cost[S']} \quad \text{and} \quad T(s, d)(s') = \Pr_{C[s](d)}(\Diamond_{S'}^{s'})$$

where  $Cost[S']$  is obtained from  $Cost$  by changing the set of goal states to  $S'$ . Note that the definition is correct as it does not depend on the delay function apart from its value  $d$  in the initial state  $s$ .

*Example 2.* Let us illustrate the construction on the fdCTMC from Section 1. The model, depicted on the left is modified to satisfy the assumption  $S_{fd} = S^{reset} \uplus S^{keep}$ : we duplicate the state *init* into another state *one*  $\in S^{keep}$ ; the states from  $S^{reset}$  are then depicted in the top row. As in Section 1, we assign cost rate 1 to all states and impulse cost 3 to every fixed-delay transition (and zero to exp-delay transitions).



On the right, there is an excerpt of the DTMDP  $\mathcal{M}$  and of the cost function  $\mathbb{E}$ . For each non-goal state, we depict only one action out of uncountably many: for state *two* it is action 0.1 with cost  $\approx 2.9$ , for state *init* it is action 0.4 with cost  $\approx 2.7$ . The costs are computed in PRISM.  $\square$

Note that there is a one-to-one correspondence between the delay functions in  $C$  and strategies in  $\mathcal{M}$ . Thus we use  $\mathbf{d}, \mathbf{d}', \dots$  to denote strategies in  $\mathcal{M}$ . Finally, let us state correctness of the reduction.

**Proposition 2.** *For any delay function  $\mathbf{d}$  it holds  $E_{C(\mathbf{d})} = E_{\mathcal{M}(\mathbf{d})}$ . Hence,*

$$Val[C] = Val[\mathcal{M}].$$

In particular, in order to solve the optimization problem for  $C$  it suffices to find an  $\varepsilon$ -optimal strategy (i.e., a delay function)  $\mathbf{d}$  in  $\mathcal{M}$ .

### 3.2 Discretization of the Uncountable MDP $\mathcal{M}$

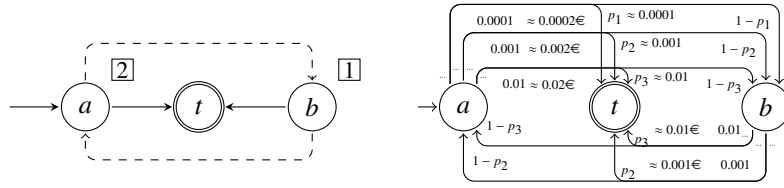
Since the MDP  $\mathcal{M}$  has uncountably many actions, it is not directly suitable for algorithmic solutions. We proceed in two steps. In the first and technically demanding step, we show that we can restrict to actions on a finite mesh. Second, we argue that we can also approximate all transition probabilities and costs by rational numbers of small bit length.



**Restricting to a Finite Mesh.** For positive reals  $\delta > 0$  and  $\bar{d} > 0$ , we define a subset of delay functions  $D(\delta, \bar{d}) = \{\mathbf{d} \mid \forall s \in S' \exists k \in \mathbb{N} : \mathbf{d}(s) = k\delta \leq \bar{d}\}$ . Here, all delays are multiples of  $\delta$  bounded by  $\bar{d}$ .

We need to argue that for the fixed  $\varepsilon > 0$  there are some appropriate values  $\delta$  and  $\bar{d}$  such that  $D(\delta, \bar{d})$  contains an  $\varepsilon$ -optimal delay function. A naïve approach would be to take any, say  $\varepsilon/2$ -optimal delay function  $\mathbf{d}$ , round it to closest delay function  $\mathbf{d}^* \in D(\delta, \bar{d})$  on the mesh, and show that the expected costs of these two functions do not differ by more than  $\varepsilon/2$ . However, this approach does not work as shown by the following example.

*Example 3.* Let us fix the fdCTMC structure  $C$  on the left (with cost rates in small boxes and zero impulse costs). An excerpt of  $\mathcal{M}$  and  $\mathbb{E}$  is shown on the right (where only a few actions are depicted).



First, we point out that  $Val[C] = 1$  as one can make sure that nearly all time before reaching  $t$  is spent in the state  $b$  that has a lower cost rate 1. Indeed, this is achieved by setting a very short delay in  $a$  and a long delay in  $b$ .

We claim that for any  $\delta > 0$  there is a near-optimal delay function  $\mathbf{d}$  such that rounding its components to the nearest integer multiples of  $\delta$  yields a large error independent of  $\delta$ . Indeed, it suffices to take a function  $\mathbf{d}$  with  $\mathbf{d}(b) = \delta$  and  $\mathbf{d}(a)$  an arbitrary number significantly smaller than  $\mathbf{d}(b)$ , say  $\mathbf{d}(a) = 0.01 \cdot \mathbf{d}(b)$ . The error produced by the rounding can then be close to 0.5. For instance, given  $\delta = 0.01$  we take a function with  $\mathbf{d}(a) = 0.0001$  and  $\mathbf{d}(b) = 0.01$ , whose rounding to the closest delay function on the finite mesh yields a constant function  $\mathbf{d}^* = (0.01, 0.01)$ . Then  $E_{C(\mathbf{d})} \approx 1.01$  and  $E_{C(\mathbf{d}^*)} \approx 1.5$ , even though the rounding does not change any transition probability or cost by more than 0.02!

The reason why the delay function  $\mathbf{d}$  is so sensitive to small perturbations is that it makes a very large number of steps before reaching  $t$  (around 200 on average) and thus the small one-step errors caused by a perturbation accumulate into a large global error. The number of steps of an  $\varepsilon$ -optimal delay functions is not bounded, in general. By multiplying both  $\mathbf{d}(a)$  and  $\mathbf{d}(b)$  by the same infinitesimally small factors we obtain an  $\varepsilon$ -optimal delay functions that make an arbitrarily high expected number of steps before reaching  $t$ .  $\square$

A crucial observation is that we do not have to show that the “naïve” rounding works for *every* near-optimal delay function. To prove that  $D(\delta, \bar{d})$  contains an  $\varepsilon$ -optimal function, it suffices to show that there is *some*  $\varepsilon/2$ -optimal function whose rounding yields error at most  $\varepsilon/2$ . Proving the existence of such well-behaved functions forms the technical core of our discretization process which is formalized below.

We start by formalizing the concept of “small perturbations”. We say that a delay function  $\mathbf{d}^*$  is  $\alpha$ -bounded by a delay function  $\mathbf{d}$  if for all states  $s, t \in S'$  we have:

1.  $|T(s, \mathbf{d}^*(s))(t) - T(s, \mathbf{d}(s))(t)| \leq \alpha$  and
2.  $\mathbb{E}(s, \mathbf{d}^*(s)) - \mathbb{E}(s, \mathbf{d}(s)) \leq \alpha$ ;

and furthermore,  $T(s, \mathbf{d}(s))(t) = 0$  iff  $T(s, \mathbf{d}^*(s))(t) = 0$ , i.e. the qualitative transition structure is preserved. (Note that  $\mathbf{d}^*$  may incur much smaller one-step costs than  $\mathbf{d}$ , but not significantly higher).

Using standard techniques of numerical analysis, we express the increase in accumulated cost caused by a bounded perturbation as a function of the worst-case (among all possible initial states) expected cost and expected number of steps before reaching the target. The number of steps is essential as discussed in Example 3 and can be easily expressed by a cost function  $\#$  that assigns 1 to every action in every state. To express the worst-case expectation of some cost function  $\$$ , we denote  $\text{Bound}[\$, \mathbf{d}] := \max_{s \in S'} E_{\mathcal{M}[s]}^{\$}(\mathbf{d})$ .

**Lemma 2.** *Let  $\alpha \in [0, 1]$  and let,  $\mathbf{d}'$  be a delay function that is  $\alpha$ -bounded by another delay function  $\mathbf{d}$ . If  $\alpha \leq \frac{1}{2 \cdot \text{Bound}[\$, \mathbf{d}] \cdot |S'|}$ , then*

$$E_{\mathcal{M}(\mathbf{d}')} \leq E_{\mathcal{M}(\mathbf{d})} + 2 \cdot \alpha \cdot \text{Bound}[\#, \mathbf{d}] \cdot (1 + \text{Bound}[\$, \mathbf{d}] \cdot |S'|).$$

The next lemma shows how to set the parameters  $\delta$  and  $\bar{d}$  to make the finite mesh  $D(\delta, \bar{d})$  “dense” enough, i.e. to ensure that for any  $\mathbf{d}$ ,  $D(\delta, \bar{d})$  contains a delay function that is  $\alpha$ -bounded by  $\mathbf{d}$ .

**Lemma 3.** *There are positive numbers  $D_1, D_2 \in \exp(\|C\|^{O(1)})$  computable in time polynomial in  $\|C\|$  such that the following holds for any  $\alpha \in [0, 1]$  and any delay function  $\mathbf{d}$ : If we put*

$$\delta := \alpha / D_1 \quad \text{and} \quad \bar{d} := |\log(\alpha)| \cdot D_2 \cdot \text{Bound}[\$, \mathbf{d}],$$

*then  $D(\delta, \bar{d})$  contains a delay function which is  $\alpha$ -bounded by  $\mathbf{d}$ .*

*Proof (Sketch).* Computing the value of  $\delta$  is easy as the derivatives of the probabilities and costs are bounded from above by the rate  $\lambda$  and the maximal cost rate, respectively. For  $\bar{d}$  we need additional technical observations, see [7] for further details.

Unfortunately, as shown in Example 3, the value  $\text{Bound}[\#, \mathbf{d}]$  can be arbitrarily high, even for near-optimal functions  $\mathbf{d}$ . Hence, we cannot use Lemma 2 right away to show that a delay function in  $D(\delta, \bar{d})$  that is  $\alpha$ -bounded by some near-optimal  $\mathbf{d}$  is also near-optimal. The crucial insight is that for any  $\varepsilon' > 0$  there are (globally)  $\varepsilon'$ -optimal delay functions that use number of steps that is *proportional* to their expected cost.

**Lemma 4.** *There is a positive number  $N \in \exp(\|C\|^{O(1)})$  computable in time polynomial in  $\|C\|$  such that the following holds: for any  $\varepsilon' > 0$ , there is a globally  $\varepsilon'/2$ -optimal delay function  $\mathbf{d}'$  with*

$$\text{Bound}[\#, \mathbf{d}'] \leq \frac{\text{Bound}[\$, \mathbf{d}']}{\varepsilon'} \cdot N. \quad (1)$$

*Proof (Sketch).* After proving the existence of globally near-optimal strategies, we suitably define the number  $N$  and take an arbitrary globally  $\varepsilon''$ -optimal delay function  $\mathbf{d}''$ , where  $\varepsilon'' \ll \varepsilon'$ . If this function *does not* satisfy (1), we conclude that it must induce the following pathological behaviour in  $C$ : the system stays for a long time in a component of its state space such that a) fixed-delay transitions are active in each state of the component, each such transition within the component having zero impulse cost; and b) function  $\mathbf{d}''$  assigns very small (in a well-defined sense) delays to all states of the component. We call such a component a *bad sink*. Intuitively, inside a bad sink the system rapidly performs one

fixed-delay transition after another, incurring only a tiny cost between two successive transitions. This allows the delay function to perform many steps while staying  $\varepsilon''$ -optimal. (In Example 3,  $\{a, b\}$  would be a bad sink for  $\mathbf{d}$ , as with high probability the cycle on these two states is completed every 0.0101 units of time, with cost 0.0102 incurred per cycle.)

To obtain a globally  $\varepsilon'$ -optimal delay function satisfying (1), we carefully modify  $\mathbf{d}''$  so as to remove all bad sinks. This is done by selecting a suitable state in each bad sink and “inflating” its delay to a sufficiently high threshold. Choosing the right state and threshold is a rather delicate process, since an improper choice might significantly increase the incurred cost. Also note that Lemma 2 cannot be used to bound the increase in cost caused by the modification, as we do not know the value of  $\text{Bound}[\#, \mathbf{d}'']$ . Instead, we utilize non-trivial insights into the structure of  $C$  and  $\mathcal{M}$ .  $\square$

By using these proportional delay functions, we reduce the perturbation error of Lemma 2 only to a function of  $\text{Bound}[\epsilon, \mathbf{d}]$ . Combining this with Lemma 3, we obtain that the delay functions in  $D(\delta, \bar{d})$  approximate all the proportional delay functions  $\mathbf{d}$  of Lemma 4, and thus  $\text{Val}[\mathcal{M}, D(\delta, \bar{d})]$  approximates  $\text{Val}[\mathcal{M}]$ . The parameters  $\delta, \bar{d}$  depend on  $\varepsilon$  and  $\text{Bound}[\epsilon, \mathbf{d}]$  of any such  $\mathbf{d}$  from Lemma 4. As these delay functions are globally  $\varepsilon$ -optimal, all such  $\text{Bound}[\epsilon, \mathbf{d}]$  can be  $\varepsilon$ -approximated by  $\overline{\text{Val}}[\mathcal{M}] := \max_{s \in S'} \text{Val}[\mathcal{M}[s]]$ .

**Proposition 3.** *For  $N$  from Lemma 4,  $D_1$  and  $D_2$  from Lemma 3, it holds that*

$$\left| \text{Val}[\mathcal{M}] - \text{Val}[\mathcal{M}, D(\delta, \bar{d})] \right| \leq \frac{\varepsilon}{2}$$

$$\text{where } \delta := \frac{\alpha}{D_1}, \bar{d} := |\log(\alpha)| \cdot D_2 \cdot (\overline{\text{Val}}[\mathcal{M}] + \varepsilon), \alpha := \frac{\varepsilon^2}{64N \cdot |S'| \cdot (1 + \overline{\text{Val}}[\mathcal{M}])^2}.$$

**Bounding  $\overline{\text{Val}}[\mathcal{M}]$**  In Proposition 3, the allowed perturbation  $\alpha$  and hence the fineness of the mesh  $\delta$  needed to obtain the required precision depend on the bound  $\overline{\text{Val}}[\mathcal{M}]$ . We first provide the following theoretical worst-case bound.

**Lemma 5.** *There is a number  $M \in \exp(\|C\|^{O(1)})$  computable in time polynomial in  $\|C\|$  such that  $\overline{\text{Val}}[\mathcal{M}] \leq M$ .*

In practice, one can obtain better bounds by computing  $\max_{s \in S} E_{C[s](\mathbf{d})}$  for an arbitrary  $\mathbf{d}$  as  $\max_{s \in S} E_{C[s](\mathbf{d})} \geq \max_{s \in S} \inf_{\mathbf{d}'} E_{C[s](\mathbf{d}')} = \overline{\text{Val}}[\mathcal{M}]$ . One can set  $\mathbf{d}$  by some heuristics (e.g. to the constant function  $1/\lambda$ ) or randomly. One can even use the minimum from a series of such computations. We believe that in most cases, this yields a significant improvement. For instance, for the 3-state model from Section 1, we get a bound  $\max_{s \in S} E_{C[s](1/\lambda)} \approx 4.3$  instead of the theoretical bound  $\overline{\text{Val}}[\mathcal{M}] \approx 55000$ .

**Representing the Finite Mesh** Since one-step costs and probabilities produced by delay functions in  $D(\delta, \bar{d})$  may be irrational, we need to approximate them by rational numbers. So let us fix  $\delta$  and  $\bar{d}$  from Proposition 3. For any  $\kappa > 0$  we define DTMDP  $\mathcal{M}_\kappa = (S', \text{Act}_\kappa, T_\kappa, s_{\text{in}}, G)$  with a cost function  $\epsilon_\kappa$  where

- the strategies are exactly delay functions from  $D(\delta, \bar{d})$ , i.e.  $\text{Act}_\kappa = \{k\delta \mid k \in \mathbb{N}, \delta \leq k\delta \leq \bar{d}\} \cup \{\infty\}$  where again  $\infty$  is enabled in  $s \in S' \setminus S_{\text{fd}}$  and the rest is enabled in  $s \in S_{\text{fd}}$ ; and

- for all  $(s, \mathbf{d}) \in S' \times \text{Act}_\kappa$  the transition probabilities in  $T_\kappa(s, \mathbf{d})$  and costs in  $\mathbb{E}_\kappa(s, \mathbf{d})$  are obtained by rounding the corresponding numbers in  $T(s, \mathbf{d})$  and  $\mathbb{E}(s, \mathbf{d})$  up to the closest multiple of  $\kappa$ .<sup>4</sup> For more details and pseudo-codes see Appendix A.

**Proposition 4.** *Let  $\varepsilon > 0$  and fix  $\kappa = (\varepsilon \cdot \delta \cdot \min R) / (2 \cdot |S'| \cdot (1 + \overline{\text{Val}}[\mathcal{M}])^2)$ , where  $\min R$  is a minimal cost rate in  $C$ . Then it holds*

$$\left| \text{Val}[\mathcal{M}, D(\delta, \bar{d})] - \text{Val}[\mathcal{M}_\kappa] \right| \leq \frac{\varepsilon}{2}.$$

*Proof (Sketch).* We use similar technique as in Lemma 2, taking advantage of the fact that probabilities and costs of each action are changed by at most  $\kappa$  by the rounding.  $\square$

**The Algorithm for Theorem 1** First the discretization step  $\delta$ , maximal delay  $\bar{d}$ , and rounding error  $\kappa$  are computed. Then the discretized DTMDP  $\mathcal{M}_\kappa$  is constructed according to the above-mentioned finite mesh representation. Finally the globally optimal delay function from  $\mathcal{M}_\kappa$  is chosen using standard polynomial algorithms for finite MDPs [34, 14]. From Propositions 3 and 4 it follows that this delay function is  $\varepsilon$ -optimal in  $\mathcal{M}$ , and thus also in  $C$  (Proposition 2).

The size of  $\mathcal{M}_\kappa$  (and its construction time) can be stated in terms of a polynomial in  $\|C\|$ ,  $\overline{\text{Val}}[\mathcal{M}]$ ,  $1/\delta$ ,  $\bar{d}$ , and  $1/\kappa$ . Examining the definitions of these parameters in Propositions 3 and 4, as well as the bound on  $\overline{\text{Val}}[\mathcal{M}]$  from Lemma 5, we conclude that the size of  $\mathcal{M}_\kappa$  and the overall running time of our algorithm are exponential in  $\|C\|$  and polynomial in  $1/\varepsilon$ . The pseudo-code of the whole algorithm is given in Appendix A.

## 4 Bounded Optimization Under Partial Observation

In this section, we address the cost optimization problem for delay functions chosen under partial observation. For an equivalence relation  $\equiv$  on  $S_{\text{fd}}$  specifying observations, and  $\underline{d}, \bar{d} > 0$ , we define  $D(\underline{d}, \bar{d}, \equiv) = \{\mathbf{d} \mid \forall s, s' : \underline{d} \leq \mathbf{d}(s) \leq \bar{d}, s \equiv s' \Rightarrow \mathbf{d}(s) = \mathbf{d}(s')\}$ .

**Theorem 2.** *There is an algorithm that for a fdCTMC structure  $C$ , a cost structure  $\text{Cost}$  with  $\mathcal{R}(s) > 0$  for all  $s \in S$ , an equivalence relation  $\equiv$  on  $S_{\text{fd}}$ ,  $\underline{d}, \bar{d} > 0$ , and  $\varepsilon > 0$  computes in time exponential in  $\|C\|$ ,  $\|\underline{d}\|$ , and  $\bar{d}$  a delay function  $\mathbf{d}$  such that*

$$\left| \inf_{\mathbf{d}' \in D(\underline{d}, \bar{d}, \equiv)} E_C(\mathbf{d}') - E_C(\mathbf{d}) \right| < \varepsilon.$$

Also, one cannot hope for polynomial complexity as the corresponding threshold problem is NP-hard, even if we restrict to instances where  $\bar{d}$  is of magnitude polynomial in  $\|C\|$ .

**Theorem 3.** *For a fdCTMC structure  $C$ , a cost structure  $\text{Cost}$  with  $\mathcal{R}(s) > 0$  for all  $s \in S$ , an equivalence relation  $\equiv$  on  $S_{\text{fd}}$ ,  $\underline{d}, \bar{d} > 0$ ,  $\varepsilon > 0$ , and  $x \in \mathbb{R}_{\geq 0}$ , it is NP-hard to decide*

$$\text{whether} \quad \inf_{\mathbf{d} \in D(\underline{d}, \bar{d}, \equiv)} E_C(\mathbf{d}) > x + \varepsilon \quad \text{or} \quad \inf_{\mathbf{d} \in D(\underline{d}, \bar{d}, \equiv)} E_C(\mathbf{d}) < x - \varepsilon$$

*(if the optimal cost lies in the interval  $[x - \varepsilon, x + \varepsilon]$ , an arbitrary answer may be given). The problem remains NP-hard even if  $\bar{d}$  is given in unary encoding.*

<sup>4</sup> More precisely, all but the largest probability in  $T(s, \mathbf{d})$  are rounded up, the largest probability is suitably rounded down so that the resulting vector adds up to 1.

For  $\bar{d}$  given in unary we get a matching upper bound.

**Theorem 4.** *The approximate threshold problem of Theorem 3 is in NP, provided that  $\bar{d}$  is given in unary.*

We leave the task of settling the exact complexity of the general problem (where  $\bar{d}$  is given in binary) to future work.

For the rest of this section we fix a fdCTMC structure  $C = (S, \lambda, P, S_{fd}, F, s_{in})$ , a cost structure  $Cost = (G, \mathcal{R}, I_P, I_F)$ ,  $\varepsilon > 0$ , and an equivalence relation  $\equiv$  on  $S_{fd}$ ,  $\underline{d}, \bar{d} > 0$ . We simply write  $D$  instead of  $D(\underline{d}, \bar{d}, \equiv)$  and again assume that  $Val[C, D] < \infty$ .

#### 4.1 Approximation Algorithm

In this Section, we address Theorem 2. First observe, that the MDP  $\mathcal{M}$  introduced in Section 3 can be due to Proposition 2 also applied in the bounded partial-observation setting. Indeed,  $E_{C(\mathbf{d})} = E_{\mathcal{M}(\mathbf{d})}$  for each  $\mathbf{d} \in D$  and thus,  $Val[C, D] = Val[\mathcal{M}, D]$  (where analogously  $Val[\mathcal{M}, D]$  denotes  $\inf_{\mathbf{d} \in D} E_{\mathcal{M}(\mathbf{d})}$ ). Furthermore, by fixing a mesh  $\delta$  and a round-off error  $\kappa$ , we define a finite DTMDP  $\mathcal{M}_D^*$  where

- actions are restricted to a finite mesh of multiples of  $\delta$  within the bounds  $\underline{d}$  and  $\bar{d}$ ; and
- probabilities and costs are rounded to multiples of  $\kappa$  as in Section 3.

To show that  $\mathcal{M}_D^*$  suitably approximates  $\mathcal{M}$  we use similar techniques as in Section 3. However, thanks to the constraints  $\underline{d}$  and  $\bar{d}$  we can show that for *every* delay function  $\mathbf{d} \in D$  the values  $Bound[\#, \mathbf{d}]$  and  $Bound[\epsilon, \mathbf{d}]$ , which feature in Lemma 2, are bounded by a function of  $\|C\|$ ,  $\underline{d}$  and  $\bar{d}$  (in particular, the bound is independent of  $\mathbf{d}$ ). This substantially simplifies the analysis. We state just the final result.

**Proposition 5.** *There is a number  $B \in \exp((\|C\| \cdot \|\underline{d}\| \cdot \bar{d})^{O(1)})$  such that for  $\delta = \varepsilon/B$  and  $\kappa = (\varepsilon \cdot \delta)/B$  it holds  $|Val[\mathcal{M}, D] - Val[\mathcal{M}_D^*]| < \varepsilon$ .*

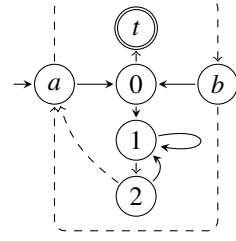
The proof of Theorem 2 is finished by the following algorithm.

- For  $\delta$  and  $\kappa$  from Proposition 5, the algorithm first constructs (in the same fashion as in Section 3) in 2-exponential time the MDP  $\mathcal{M}_D^*$ .
- Then it finds an optimal strategy  $\mathbf{d}$  (which also satisfies  $|E_{C(\mathbf{d})} - \inf_{\mathbf{d}'} E_{C(\mathbf{d}')}| < \varepsilon$ ) by computing  $E_{\mathcal{M}_D^*(\mathbf{d})}$  for every (MD) strategy  $\mathbf{d}$  of  $\mathcal{M}_D^*$  in the set  $D$ .

The algorithm runs in 2-EXPTIME because there are  $\leq |Act_\varepsilon|^{|S|}$  strategies which is exponential in  $\|C\|$ ,  $\|\underline{d}\|$ , and  $\bar{d}$  as  $|Act_\varepsilon|$  is exponential in these parameters. The correctness follows from Propositions 2, 5, proving Theorem 2.

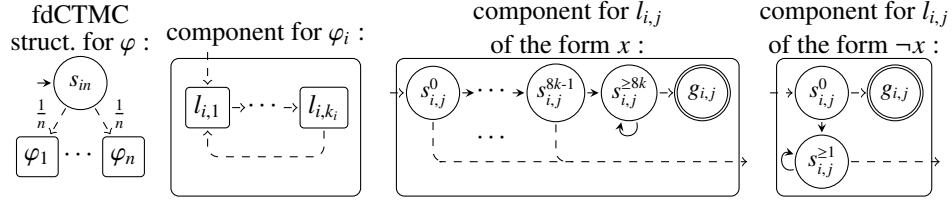
**Challenges of Unbounded Optimization** The proof of Proposition 5 is simpler than the techniques from Section 3 because we work with the compact space bounded by  $\underline{d}$  and  $\bar{d}$ . This restriction is not easy to lift; the techniques from Section 3 cannot be easily adapted to *unbounded* optimization under partial observation.

The reason is that local adaptation of the delay function (heavily applied in the proof of Lemma 4) is not possible as the delays are not independent. Consider on the right a variant of Example 3 with components  $a$  and  $b$  being switched by fixed-delay transitions. All states have cost rate 1 and all transitions have cost 0; furthermore, all states are in one class of equivalence of  $\equiv$ . If in state  $a$  or  $b$  more than one exp-delay transition is taken before a fixed-delay transition, a long detour via state 1 is taken. In order to avoid it and to optimize the cost, one needs to set the one common delay as close as possible to 0. Contrarily, in order to decrease the expected number of visits from  $a$  to  $b$  before reaching  $t$  which is crucial for the error bound, one needs to increase the delay.



## 4.2 Complexity of the Threshold Problem

We now turn our attention to Theorem 3. We show the hardness by reduction from SAT. Let  $\varphi = \varphi_1 \wedge \dots \wedge \varphi_n$  be a propositional formula in conjunctive normal form (CNF) with  $\varphi_i = (l_{i,1} \vee \dots \vee l_{i,k_i})$  for each  $1 \leq i \leq n$  and with the total number of literals  $k = \sum_{i=1}^n k_i$ . As depicted in the following figure, the fdCTMC structure  $C_\varphi$  is composed of  $n$  components (one per clause), depicted by rectangles. The component of each clause is formed by a cycle of sub-components (one per literal) connected by fixed-delay transitions. Positive literals are modelled differently from negative literals.



The cost structure  $Cost_\varphi$  assigns rate cost 1 to every state, and impulse cost 0 to every transition; the goal states are depicted by double circles and exp-delay transitions are depicted with heavier heads to distinguish from the dashed fixed-delay transitions. We require  $s_{i,j}^0 \equiv s_{i',j'}^0$  iff the literals  $l_{i,j}$  and  $l_{i',j'}$  have the same variable. Furthermore, let  $D$  denote  $D(0.01, 16k, \equiv)$ . Note that  $\bar{d} = 16k$  is linear in  $\|\varphi\|$  and thus it can be encoded in unary. We obtain the following:

**Proposition 6.** *For a formula  $\varphi$  in CNF with  $k$  literals,  $C_\varphi$  and  $Cost_\varphi$  are constructed in time polynomial in  $k$  and, furthermore,*

$$Val[C_\varphi, D] < 17k^2 \text{ if } \varphi \text{ is satisfiable} \quad \text{and} \quad Val[C_\varphi, D] > 17k^2 + 1, \text{ otherwise.}$$

*Proof (Sketch).* If  $\varphi$  is satisfiable, let  $\nu$  denote the satisfying truth assignment. We set  $\mathbf{d}(s_{in}) := \underline{d}$  and  $\mathbf{d}(s_{i,j}^0) := \bar{d}$  if  $\nu(X) = 1$  and  $\mathbf{d}(s_{i,j}^0) := \underline{d}$  if  $\nu(X) = 0$ , where  $X$  is the variable of the literal  $l_{i,j}$  and arbitrarily in other states. In  $C_\varphi(\mathbf{d})$  in the component for any TRUE literal  $l_{i,j}$ , i.e. with  $\nu(l_{i,1}) = 1$ , the goal is reached from  $s_{i,j}^0$  with probability  $> 0.99$  before leaving the component. Indeed, the probability to take no exponential transition within time

0.01 is  $> 0.99$  and the probability to take at least  $8k$  exponential transitions within time  $16k$  is  $> 0.99$  for any  $k \in \mathbb{N}$ . As each clause  $\varphi_i$  has at most  $k$  literals and at least one TRUE literal, the expected cost incurred in the component for  $\varphi_i$  is at most  $(16k \cdot k)/0.99 < 17k^2$ .

The other implication is based on the observation that there is no delay length guaranteeing high probability of going directly into the target from both components for a positive and a negative literal.  $\square$

The reduction proves NP-hardness as it remains to set  $x := 2k^2 + \frac{1}{2}$  and  $\varepsilon := \frac{1}{2}$ .

**NP Membership for Unary  $\bar{d}$**  To prove Theorem 4 we give an algorithm which, for a given approximate threshold  $x > 0$ , consists of

- *first* guessing the delay function  $\mathbf{d}$  of  $\mathcal{M}_D^*$  that is in the set  $D$  such that  $E_{\mathcal{M}_D^*(\mathbf{d})} < x$ ,
- *then* constructing just the fragment  $\mathcal{M}_\mathbf{d}$  of  $\mathcal{M}_D^*$  used by the guessed function  $\mathbf{d}$ . Here  $\mathcal{M}_\mathbf{d} = (S', \{\infty\}, T_\mathbf{d}, G, \mathbb{E}_\mathbf{d})$  where the transition probabilities and costs coincide with  $\mathcal{M}$  for states in  $S' \setminus S_{\text{fd}}$  and in any state  $s \in S_{\text{fd}}$  are defined by  $T_\mathbf{d}(s, \infty) = T^*(s, \mathbf{d}(s))$  and  $\mathbb{E}_\mathbf{d}(s, \infty) = \mathbb{E}^*(s, \mathbf{d}(s))$  (here  $T^*(s, \mathbf{d}(s))$  and  $\mathbb{E}^*(s, \mathbf{d}(s))$  are as in  $\mathcal{M}_D^*$ ).
- Last, for  $\sigma : s \mapsto \infty$ , the algorithm computes  $y = E_{\mathcal{M}_\mathbf{d}(\sigma)}$  by standard methods and *accepts* iff  $y < x$ .

Note that when  $\bar{d}$  is encoded in unary, both  $\mathbf{d}$  and  $\mathcal{M}_\mathbf{d}$  are of bit size that is polynomial in the size of the input. Hence,  $\mathbf{d}$  and  $\mathcal{M}_\mathbf{d}$  can be constructed in non-deterministic polynomial time (although the whole  $\mathcal{M}_D^*$  is of exponential size in this unary case). The expected total cost  $x$  in  $\mathcal{M}_\mathbf{d}(\sigma)$  that has polynomial size can be also computed in polynomial time. The correctness of the algorithm easily follows from Proposition 5; for an explicit proof see Appendix D.3.

## 5 Conclusions

In this paper, we introduced the problem of synthesizing timeouts for fixed-delay CTMC. We study two variants of this problem, show that they are effectively solvable, and obtain provable worst-case complexity bounds. First, for *unconstrained optimization*, we present an approximation algorithm based on a reduction to a discrete-time Markov decision process and a standard optimization algorithm for this model. Second, we approximate the case of *bounded optimization under partial observation* also by a MDP. However, a restriction of the class of strategies twists it basically into a partial-observation MDP (where only memoryless deterministic strategies are considered). We give a 2-exponential approximation algorithm (which becomes exponential if one of the constraints is given in unary) and show that the corresponding decision problem is NP-hard.

The correctness of our algorithms stems from non-trivial insights into the behaviour of fdCTMC that we deem to be interesting in their own right. Hence, we believe that techniques presented in this paper lay the ground for further development of performance optimization via timeout synthesis.

## References

1. R. Alur, C. Courcoubetis, and D.L. Dill. Model-checking for probabilistic real-time systems. In *Proceedings of ICALP'91*, volume 510 of *LNCS*, pages 115–136. Springer, 1991.

2. R. Alur, T.A. Henzinger, and M.Y. Vardi. Parametric real-time reasoning. In *STOC*, pages 592–601. ACM, 1993.
3. N.C. Audsley and A. Grigg. Timing analysis of the ARINC 629 databus for real-time applications. *Microprocessors and Microsystems*, 21(1):55–61, 1997.
4. C. Baier, N. Bertrand, P. Bouyer, T. Brihaye, and M. Größer. Probabilistic and topological semantics for timed automata. In *FSTTCS*, volume 4855 of *LNCS*, pages 179–191. Springer, 2007.
5. C. Baier, N. Bertrand, P. Bouyer, T. Brihaye, and M. Größer. Almost-sure model checking of infinite paths in one-clock timed automata. In *LICS*, pages 217–226. IEEE, 2008.
6. T. Brázdil, V. Forejt, J. Krčál, J. Kretínský, and A. Kucera. Continuous-time stochastic games with time-bounded reachability. *Inf. Comput.*, 224:46–70, 2013.
7. T. Brázdil, L. Korenčiak, J. Krčál, P. Novotný, and V. Řehák. Optimizing performance of continuous-time stochastic systems using timeout synthesis. *CoRR*, abs/1407.4777, 2014.
8. T. Brázdil, J. Krčál, J. Křetínský, A. Kučera, and V. Řehák. Stochastic real-time games with qualitative timed automata objectives. In *CONCUR*, volume 6269 of *LNCS*, pages 207–221. Springer, 2010.
9. T. Brázdil, J. Krčál, J. Křetínský, and V. Řehák. Fixed-delay events in generalized semi-markov processes revisited. In *CONCUR*, volume 6901 of *LNCS*, pages 140–155. Springer, 2011.
10. P. Buchholz, E.M. Hahn, H. Hermanns, and L. Zhang. Model checking algorithms for CTMDPs. In *CAV*, volume 6806 of *LNCS*, pages 225–242. Springer, 2011.
11. L. Carnevali, L. Ridi, and E. Vicario. A quantitative approach to input generation in real-time testing of stochastic systems. *IEEE Trans. Softw. Eng.*, 39(3):292–304, 2013.
12. T. Chen, T. Han, J.-P. Katoen, and A. Mereacre. Quantitative model checking of continuous-time Markov chains against timed automata specifications. In *LICS*, pages 309–318. IEEE, 2009.
13. H. Choi, V. G. Kulkarni, and K. S. Trivedi. Transient analysis of deterministic and stochastic petri nets. In *Application and Theory of Petri Nets*, pages 166–185, 1993.
14. K. Etessami, D. Wojtczak, and M. Yannakakis. Recursive stochastic games with positive rewards. In *Proceedings of ICALP 2008, Part I*, volume 5125 of *LNCS*, pages 711–723. Springer, 2008.
15. W. Feller. *An Introduction to Probability Theory and Its Applications, Vol. I*. Wiley, 1968.
16. C. Guet, A. Gupta, T. Henzinger, M. Mateescu, and A. Sezgin. Delayed continuous-time Markov chains for genetic regulatory circuits. In *CAV*, volume 7358 of *LNCS*, pages 294–309. Springer, 2012.
17. C. Haase, S. Kreutzer, J. Ouaknine, and J. Worrell. Reachability in succinct and parametric one-counter automata. In *CONCUR*, volume 5710 of *LNCS*, pages 369–383. Springer, 2009.
18. E. M. Hahn, H. Hermanns, and L. Zhang. Probabilistic reachability for parametric markov models. *STTT*, 13(1):3–19, 2011.
19. T. Han, J.P. Katoen, and A. Mereacre. Approximate parameter synthesis for probabilistic time-bounded reachability. In *Real-Time Systems Symposium*, pages 173–182. IEEE, 2008.
20. E. Isaacson and H.B. Keller. *Analysis of Numerical Methods*. Wiley, 1966.
21. A. Jensen. Markov chains as an aid in the study of Markov processes. *Skand. Aktuarietidskr.*, 36:87–91, 1953.
22. P.G. Jensen and J.H. Taankvist. Learning optimal scheduling for time uncertain settings. Student project, Aalborg University, 2014.
23. S.K. Jha and C.J. Langmead. Synthesis and infeasibility analysis for stochastic models of biochemical systems using statistical model checking and abstraction refinement. *TCS*, 412(21):2162–2187, 2011.
24. J.G. Kemeny and J.L. Snell. *Finite Markov chains*. D. Van Nostrand Company, 1960.
25. M. Khaksari and C. Fischione. Performance analysis and optimization of the joining protocol for a platoon of vehicles. In *ISCCSP*, pages 1–6. IEEE, 2012.
26. L. Korenčiak, J. Krčál, and V. Řehák. Dealing with zero density using piecewise phase-type approximation. In *EPEW*, volume 8721 of *LNCS*, pages 119–134. Springer, 2014.



27. M. Kwiatkowska, G. Norman, R. Segala, and J. Sproston. Verifying quantitative properties of continuous probabilistic timed automata. In *CONCUR*, volume 1877 of *LNCS*, pages 123–137. Springer, 2000.
28. Ch. Lindemann. An improved numerical algorithm for calculating steady-state solutions of deterministic and stochastic petri net models. *Perform. Eval.*, 18(1):79–95, 1993.
29. M. A. Marsan and G. Chiola. On Petri nets with deterministic and exponentially distributed firing times. In *Advances in Petri Nets*, pages 132–145. Springer, 1987.
30. M.R. Neuhäusser and L. Zhang. Time-bounded reachability probabilities in continuous-time Markov decision processes. In *QEST*, pages 209–218. IEEE, 2010.
31. M.F. Neuts. *Matrix-geometric Solutions in Stochastic Models: An Algorithmic Approach*. Courier Dover Publications, 1981.
32. J.R. Norris. *Markov Chains*. Cambridge University Press, 1998.
33. R. Obermaisser. *Time-Triggered Communication*. CRC Press, 1st edition, 2011.
34. M.L. Puterman. *Markov Decision Processes*. Wiley, 1994.
35. K. Ramamritham and J.A. Stankovic. Scheduling algorithms and operating systems support for real-time systems. *Proceedings of the IEEE*, 82(1):55–67, 1994.
36. K.B. Tiassou. *Aircraft operational reliability – A Model-based approach and case studies*. PhD thesis, Université de Toulouse, 2013.
37. N. Wolovick, P. R. D’Argenio, and H. Qu. Optimizing probabilities of real-time test case execution. In *ICST*, pages 446–455. IEEE, 2009.

## A Algorithms

In this section, we present pseudo-codes for algorithms.

---

**Procedure** Discretize( $C, Cost, \delta, \kappa, \ell, u$ )

---

**input** : fdCTMC structure  $C$ , cost structure  $Cost$ , and  $\delta, \kappa, \ell, u > 0$ .

**output** : A DTMDP  $\mathcal{M} = (S', Act, T, s_{in}, G)$  with a cost function  $\mathbb{E}$ .

- 1 Construct  $S', s_{in}$ , and  $G$  as specified in Subsection 3.1
  - 2  $Act := \{k\delta \mid k \in \mathbb{N} \text{ and } \ell < k\delta \leq u\} \cup \{\infty\}$
  - 3 **for**  $s \in S' \cap S_{fd}$  **and**  $d \in Act \setminus \{\infty\}$  **do**
  - 4     |     Compute distribution  $T(s, d)$  and cost in  $\mathbb{E}(s, d)$  using  $C[s](d)$  and  $\kappa$
  - 5 **for**  $s \in S' \setminus S_{fd}$  **do**  $T(s, \infty) := P(s, \cdot); \mathbb{E}(s, \infty) := \mathcal{R}(s) + \sum_{s' \in S'} P(s, s') \cdot \mathcal{I}_p(s, s')$
  - 6 set  $T$  to  $\perp$  and  $\mathbb{E}(s, a)$  to 0 for all remaining  $s \in S'$  and  $a \in Act$
- 

---

**Algorithm 1:** Unconstrained optimization

---

**input** : fdCTMC structure  $C$ , cost structure  $Cost$  with  $\mathcal{R}(s) > 0$  for all  $s \in S$ , and  $\varepsilon > 0$ .

**output** : The delay function  $\mathbf{d}$  that is  $\varepsilon$ -optimal in  $C$  and  $Cost$ .

- 1 Compute  $\overline{Val}[\mathcal{M}], \delta, \kappa$ , and  $\bar{d}$
  - 2 Construct  $\mathcal{M}_\kappa$  and  $\mathbb{E}_\kappa$  by procedure Discretize( $C, Cost, \delta, \kappa, \delta, \bar{d}$ )
  - 3  $\mathbf{d} :=$  optimal strategy in  $\mathcal{M}_\kappa$  and  $\mathbb{E}_\kappa$      // e.g. alg. [34] in PTIME in  $|\mathcal{M}_\kappa|$  and  $|\mathbb{E}_\kappa|$
- 

---

**Algorithm 2:** Bounded optimization under partial observation

---

**input** : fdCTMC structure  $C$ , cost structure  $Cost$  with  $\mathcal{R}(s) > 0$  for all  $s \in S$ , equivalence relation  $\equiv$  on  $S_{fd}$ ,  $\underline{d}, \bar{d} > 0$ , and  $\varepsilon > 0$ .

**output** : The delay function  $\mathbf{d}$  that is  $\varepsilon$ -optimal in  $C, Cost, \underline{d}, \bar{d}$ , and  $\equiv$ .

- 1 Compute  $\overline{Val}[\mathcal{M}], \delta$ , and  $\kappa$
  - 2 Construct  $\mathcal{M}_D^*$  and  $\mathbb{E}_D$  by procedure Discretize( $C, Cost, \delta, \kappa, \underline{d}, \bar{d}$ )
  - 3  $\mathbf{d} :=$  optimal strategy in  $\mathcal{M}_D^*, \mathbb{E}_D$ , and  $\equiv$      // standard algorithms for POMDP
- 

Let us discuss in closer detail, how the transition probabilities and costs are approximated.

**Approximating Transition Probabilities and Costs** Note that for  $s \in S' \setminus S_{fd}$  and  $d = \infty$ , the approximation is easy as these states behave as in CTMC and so the probabilities in  $T(s, d)$  and cost  $\mathbb{E}(s, d)$  are already expressed as rational numbers.

Let us fix a state  $s \in S' \cap S_{fd}$ ,  $d \in Act \setminus \{\infty\}$ , and error bound  $\kappa \in \mathbb{R}_{>0}$ . First, we compute all probabilities in  $T(s, d)$  and cost  $\mathbb{E}(s, d)$  up to an absolute error of  $\kappa/2$  employing transient analysis of  $C[s](d)$ . Then we round up (to the closest multiple of  $\kappa$ ) cost and all probabilities except the highest one, that is accordingly rounded down so that the sum of probabilities is 1.

Before the transient analysis we need to modify  $C[s](d)$  so that all states  $S' \setminus \{s\}$  are (1) absorbing w.r.t. both exponential and fixed-delay transitions and (2) incur neither rate nor impulse costs. If there is a fixed-delay transition leading to  $s$ , then we create an absorbing copy of  $s$  where we redirect all such transitions. After this transformation we obtain a (subordinated) continuous-time Markov chain  $\bar{C}$ , and  $T(s, d)$  is a transient

distribution in time  $d$  and  $\mathbb{E}(s, d)$  is expected accumulated cost up to time  $d$  in this  $\bar{C}$ . We denote the modified transition matrices by  $\bar{P}$  and  $\bar{F}$  and the modified cost structure by  $(\bar{R}, \bar{I}_P, \bar{I}_F)$ . Similarly to the method of uniformization [21] or methods in [28], it is easy to show that

**Lemma 6.** *For any state  $s \in S' \cap S_{fd}$ , and  $d \in \mathbb{R}_{>0}$ ,*

$$\begin{aligned} T(s, d) &= \sum_{i=0}^{\infty} \text{Pois}_{\lambda d}(i) \cdot \left( \mathbf{1}_s \cdot \bar{P}^i \right) \cdot \bar{F} \\ \mathbb{E}(s, d) &= \sum_{i=0}^{\infty} \text{Pois}_{\lambda d}(i) \left( \sum_{j=0}^{i-1} \left( \mathbf{1}_s \cdot \bar{P}^j \right) \cdot \left( \frac{d \cdot \bar{R}}{i+1} + \bar{J}_Q \right) + \left( \mathbf{1}_s \cdot \bar{P}^i \right) \cdot \left( \frac{d \cdot \bar{R}}{i+1} + \bar{J}_F \right) \right) \end{aligned}$$

where  $\mathbf{1}_s$  denotes the unit vector of state  $s$ ,  $\text{Pois}_{\lambda d}$  denotes the probability mass function of the Poisson distribution with parameter  $\lambda d$ , and  $\bar{J}_Q, \bar{J}_F : S \rightarrow \mathbb{R}_{\geq 0}$  assign to each state the expected impulse reward of the next exponential or fixed-delay transition, i.e.  $\bar{J}_Q(s) = \sum_{s'} \bar{P}(s, s') \cdot \bar{I}_P(s, s')$  and  $\bar{J}_F(s) = \sum_{s'} \bar{F}(s, s') \cdot \bar{I}_F(s, s')$ .

Indeed, the distribution over states after  $j$  exponential transitions is  $\mathbf{1}_s \cdot \bar{P}^j$ ; the probability that exactly  $i$  exponential transitions occur before time  $d$  is  $\text{Pois}_{\lambda d}(i)$ ; and under the condition that  $i$  transitions occur, the expected time spent in each of the  $i+1$  visited states is  $d/(i+1)$  [28].

Finally, note that for any fixed error  $\gamma > 0$ , we can easily compute a bound  $I$  such that  $\sum_{i=I}^{\infty} \text{Pois}_{\lambda d}(i) < \gamma/2$ . Hence the  $T(s, d)$  can be approximated up to  $\gamma/2$  by truncating the infinite sums at  $I$ . Similarly there is a bound  $J$  for approximation of  $\mathbb{E}(s, d)$  since  $\text{Pois}_{\lambda d}$  decreases exponentially while the costs increase linearly for increasing  $i$ .

## B Proofs for Section 2

### B.1 Definition of probability space for fdCTMC

For a run  $(s_0, d_0)t_0(s_1, d_1)t_1 \cdots$  and for each  $i \in \mathbb{N}_0$  we define function  $\text{next}(s_i, d_i; t_i)$  as follows:

- If  $t_i < d_i$  (an exp-delay transition occurs),

$$\text{next}(s_i, d_i; t_i) = \begin{cases} d_i - t_i & \text{if } s_{i+1} \in S_{fd} \text{ and } s_i \in S_{fd}, \\ \mathbf{d}(s_{i+1}) & \text{if } s_{i+1} \in S_{fd} \text{ and } s_i \notin S_{fd}, \\ \infty & \text{if } s_{i+1} \notin S_{fd}, \end{cases}$$

- If  $t_i = d_i$  (a fixed-delay transition occurs),

$$\text{next}(s_i, d_i; t_i) = \begin{cases} \mathbf{d}(s_{i+1}) & \text{if } s_{i+1} \in S_{fd}, \\ \infty & \text{if } s_{i+1} \notin S_{fd}. \end{cases}$$

We define the probability measure  $\Pr_{C(\mathbf{d})}$  over the measurable space  $(\Omega, \mathcal{F})$  where  $\Omega$  is the set of all runs initiated in  $s_{in}$  and  $\mathcal{F}$  is a  $\sigma$ -field over  $\Omega$  generated by the set all of *cylinder sets* of the form  $R_{I_0, \dots, I_{n-1}}^{s_0, \dots, s_n} = \{(s'_0, d_0)t_0 \cdots \in \Omega \mid \forall i \leq n : s'_i = s_i, \forall i < n : t_i \in I_i\}$ , here  $s_0, \dots, s_n \in S$  and  $I_0, \dots, I_{n-1} \subseteq \mathbb{R}_{\geq 0}$  are intervals. Given such a cylinder set  $R_{I_0, \dots, I_{n-1}}^{s_0, \dots, s_n}$ , we define the probability that a run of  $C(\mathbf{d})$  belongs to  $R_{I_0, \dots, I_{n-1}}^{s_0, \dots, s_n}$  by

$$\Pr_{C(\mathbf{d})}^{s_{in}} \left[ R_{I_0, \dots, I_{n-1}}^{s_0, \dots, s_n} \right] := \mathbf{1}_{s_0 = s_{in}} \cdot \int_{t_0 \in I_0} \cdots \int_{t_{n-1} \in I_{n-1}} \Pr_{\rightarrow}(s_0, d_0, s_1; dt_0) \cdots \Pr_{\rightarrow}(s_{n-1}, d_{n-1}, s_n; dt_{n-1})$$

where  $d_0 = \mathbf{d}(s_0)$  if  $s_0 \in S_{fd}$ , and  $d_0 = \infty$  otherwise;  $d_{i+1} = \text{next}(s_i, d_i, t_i)$  for each  $i \in \mathbb{N}_0$ ; and the probability measure  $\Pr_{\rightarrow}(s, d, s'; \cdot)$  stands for the probability of moving to  $s'$  in a given time interval when staying in  $(s, d)$ , i.e. for each  $s, s' \in S$ ,  $d \in \mathbb{R}_{>0} \cup \{\infty\}$ , and an interval  $[a, b] \subseteq \mathbb{R}_{\geq 0}$

$$\Pr_{\rightarrow}(s, d, s'; [a, b]) := \underbrace{\mathbf{1}_{a < d} \cdot P(s, s') \cdot (1 - e^{-\lambda(\min\{b, d\} - a)})}_{\text{exp-delay transition in the interval } [a, \min\{b, d\}]} + \underbrace{\mathbf{1}_{d \in [a, b]} \cdot F(s, s') \cdot e^{-\lambda d}}_{\text{fixed-delay transition after time } d \in [a, b]}.$$

The probability measure  $\Pr_{C(\mathbf{d})}^{s_{in}}$  then extends uniquely to all sets of  $\mathcal{F}$ .

## B.2 Proof of Proposition 1

**Proposition 1.** There is a polynomial-time algorithm that for a given fdCTMC  $C(\mathbf{d})$ , cost structure  $Cost$  with goal states  $G$ , and an approximation error  $\varepsilon > 0$  computes  $x \in \mathbb{R}_{>0} \cup \{\infty\}$  and  $p_s \in \mathbb{R}_{>0}$ , for each  $s \in G$ , such that

$$\left| E_{C(\mathbf{d})} - x \right| < \varepsilon \quad \text{and} \quad \left| \Pr_{C(\mathbf{d})}(\Diamond_G^s) - p_s \right| < \varepsilon$$

where  $\Diamond_G^s$  is the set of runs that reach  $s$  as the first state of  $G$  (after at least one transition).

*Proof.* This proposition follows directly from results presented in [28]. For details how the transition probabilities and costs of embedded DTMC are computed, please refer to Appendix A.

## B.3 Definition of probability space for DTMDP

Every strategy  $\sigma$  uniquely determines a probability space  $(Run_{\mathcal{M}}, \mathcal{F}_{\mathcal{M}}, \Pr_{\mathcal{M}(\sigma)})$ , where  $Run_{\mathcal{M}}$  is the set of all *runs*, i.e. infinite alternating sequences of vertices and actions  $v_0 a_1 v_1 a_2 v_2 \cdots \in (V \cdot Act)^\omega$ ;  $\mathcal{F}_{\mathcal{M}}$  is the sigma-field generated by all sets of runs of the form  $\{\rho \mid \rho \text{ has prefix } w\}$  for all *finite paths*  $w$ , i.e. prefixes of runs ending with a vertex; and  $\Pr_{\mathcal{M}(\sigma)}$  is the unique probability measure such that for every finite path  $w = v_0 a_1 v_1 \dots a_n v_n$  it holds

$$\Pr_{\mathcal{M}(\sigma)}^{v_{in}}(\{\rho \mid \rho \text{ has prefix } w\}) = \begin{cases} 0 & \text{if } \sigma(v_{i-1}) \neq a_i \text{ for some } 1 \leq i \leq n, \\ \prod_{i=1}^n T(v_{i-1}, a_i)(v_i) & \text{otherwise.} \end{cases}$$

## C Proofs for Section 3

### C.1 Necessary definitions

We denote  $\mathbb{N}_0$  the natural numbers with zero, i.e.  $\mathbb{N} \cup \{0\}$ .

**Definition 3.** Let  $C = (S, \lambda, P, S_{\text{id}}, F, s_{\text{in}})$  be a fdCTMC structure and  $\text{Cost} = (G, \mathcal{R}, \mathcal{I}_P, \mathcal{I}_F)$  be the cost structure over  $C$ . We define the minimal cost, maximal cost and minimal branching probability as

- $\max R = \max\{\mathcal{R}(s) \mid s \in S\} \cup \{\mathcal{I}_i(s, s') > 0 \mid s, s' \in S \text{ and } i \in \{P, F\}\}$
- $\min R = \min\{\mathcal{R}(s) \mid s \in S\} \cup \{\mathcal{I}_i(s, s') > 0 \mid s, s' \in S \text{ and } i \in \{P, F\}\}$ , and
- $\min P = \min\{P(s, s') > 0 \mid s, s' \in S\} \cup \{F(s, s') > 0 \mid s, s' \in S\}$ , respectively.

We will use these constants also for DTMDPs in this appendix. All DTMDPs in appendix were created of some given fdCTMC  $C$ . Thus  $\min R$ ,  $\max R$  and  $\min P$  are defined for all DTMDPs according the source fdCTMC.

Let  $(\Omega, \mathcal{F}, \mathcal{P})$  be arbitrary probability space. We define conditional probability  $\mathcal{P}(A, B)$  for some events  $A, B \in \mathcal{F}$  as  $\mathcal{P}(A, B) = \frac{\mathcal{P}(A \cap B)}{\mathcal{P}(B)}$ . Moreover let  $G$  be random variable. The expected value for discrete random variable  $G$  denoted as  $\mathbb{E}[G]$  is  $\mathbb{E}[G] = \sum_{g \in \text{Im}(G)} g \cdot \mathcal{P}(G = g)$ , where  $\text{Im}$  is function that returns and image of a function. Finally the conditional expected value of  $G$  conditioned by an event  $A \in \mathcal{F}$  as  $\mathbb{E}[G \mid A] = \sum_{g \in \text{Im}(G)} g \cdot \mathcal{P}(G = g \mid A)$ . Standard definition is also used to define  $\mathbb{E}$  for continuous random variables.

**Definition 4.** For a DTMDP  $\mathcal{M}$  and set  $U \subseteq S'$  we define the series of random variables  $(H_n^U)_n$ , where for each  $n \in \mathbb{N}$  the  $H_n^U : (S' \cdot \text{Act})^\omega \rightarrow \mathbb{N}_0$  gives the jump of the  $n$ th visit of  $U$  before reaching  $G$  (stopping time), i.e.

$$H_n^U(s_0 a_1 s_1 a_2 s_2 \dots) = \begin{cases} i_n & \text{if } i_n < l, \\ l & \text{if } i_n \geq l, \\ \infty & \text{if } i_n = l = \infty, \end{cases}$$

where  $l \geq 0$  is minimal such that it holds  $s_l \in G$  and  $i_n \geq 0$  is minimal such that there are  $i_1, \dots, i_{n-1}$  such that for each  $j, k \leq n$  it holds that  $s_{i_j} \in U$  and if  $j \neq k$  then  $i_j \neq i_k$ . If no such  $i_n$  or  $l$  exist we define them to be equal to  $\infty$ .

**Definition 5.** Let  $\mathcal{M}$  be a DTMDP. We define a jump chain for  $\mathcal{M}$  as a series of random variables  $(X_n^M)_n$  where for all  $n \in \mathbb{N}_0$ ,  $X_n^M : (S' \cdot \text{Act})^\omega \rightarrow S'$  and

$$X_n^M(s_0 a_1 s_1 a_2 s_2 \dots) = s_n.$$

We define variable  $X_\infty^M$  as  $X_\infty^M : (S' \cdot \text{Act})^\omega \rightarrow \{\perp\}$ .

**Definition 6.** Let  $\mathcal{M}$  be a DTMDP. For sets  $U, \overline{U} \subseteq S'$  we define random variable  $\#_U^{\overline{U}} : (S' \cdot \text{Act})^\omega \rightarrow \mathbb{N}_0$  giving each run number of visits of states from  $U$  before reaching any state from  $\overline{U}$ , i.e.

$$\#_U^{\overline{U}}(\omega) = \sum_{i=0}^{H_1^{\overline{U}}} 1_{X_i^M \in U},$$

where  $1$  is indicator function.

Let  $\mathcal{M}$  be a DTMDP and  $\mathbf{d}$  be a delay function. For each  $i \in \mathbb{N}_0$  and  $s \in S'$  (abusing notation) we define "delay function"  $\mathbf{d}[s/d/i]$  that assigns to every state  $s' \in S' \setminus \{s\}$  delay  $\mathbf{d}(s')$ , but it assigns to state  $s$  delay  $d$  until it is reached  $i$  times and delay  $\mathbf{d}(s)$  afterwards. Moreover we define  $\mathbf{d}[s/d]$  as  $\mathbf{d}[s/d/1]$ .

## C.2 Proof of Lemma 1

**Lemma 1.** For any delay functions  $\mathbf{d}, \mathbf{d}'$  we have  $E_{C(\mathbf{d})} = \infty$  if and only if  $E_{C(\mathbf{d}')} = \infty$ .

*Proof.* Note that the only way how  $E_{C(\mathbf{d})} = \infty$  for delay function  $\mathbf{d}$  is to reach bottom strongly connected component (BSCC) in  $C(\mathbf{d})$  with positive probability. Runs that reach such BSCC stay there forever and incur infinite cost. Observe that for each  $s \in S$  and any delay function  $\mathbf{d}$  it holds that  $0 < \mathbf{d}(s) < \infty$ . Thus there is positive probability of reaching BSCC in  $C(\mathbf{d})$  for some delay function  $\mathbf{d}$  if only if there is positive probability to reach exactly same BSCC in  $C(\mathbf{d}')$  also using any other delay function  $\mathbf{d}'$ .

## C.3 Proof of Proposition 2

**Proposition 2.** For any delay function  $\mathbf{d}$  it holds  $E_{C(\mathbf{d})} = E_{\mathcal{M}(\mathbf{d})}$ . Hence,

$$Val[C] = Val[\mathcal{M}].$$

*Proof.* Let us fix  $\mathbf{d}$  and let  $R$  and  $R'$  denote the set of runs in  $\mathcal{M}(\mathbf{d})$  and  $C(\mathbf{d})$ , respectively. Note that for each  $n \in \mathbb{N}$ , both  $R$  and  $R'$  can be partitioned into countable collections of sets  $\mathcal{R}_n = \{R_{s_0 \dots s_n} \mid s_0, \dots, s_n \in S\}$  and  $\mathcal{R}'_n = \{R'_{s_0 \dots s_n} \mid s_0, \dots, s_n \in S\}$  where

$$\begin{aligned} R_{s_0 \dots s_n} &:= \{s'_0 s'_1 \dots \mid \forall i \leq n : s'_i = s_i\} \\ R'_{s_0 \dots s_n} &:= \{(s'_0, d_0) t_0 \dots \mid s'_0 = s_0, \exists i_1 < \dots < i_n : \\ &\quad \forall i > 0 \text{ such that } i = i_k \text{ for some } k : s'_i = s_k, s'_i \notin S_{\text{fd}} \vee d_{i-1} = t_{i-1} \vee s'_i \in G, \\ &\quad \forall i > 0 \text{ such that } i \neq i_k \text{ for any } k : s_i \in S_{\text{fd}} \setminus G, d_{i-1} \neq t_{i-1}\} \end{aligned}$$

Intuitively, the indices  $s_0, \dots, s_n$  in the latter definition are the states where fixed-delay transitions are not active or become newly active, i.e. states where the "big steps" that are simulated by single steps in  $\mathcal{M}(\mathbf{d})$  start. For each  $n \in \mathbb{N}$ , let  $W_n$  and  $W'_n$  denote the cost accumulated in first  $n$  steps of  $\mathcal{M}(\mathbf{d})$  and first  $n$  "big steps" of  $C(\mathbf{d})$ , respectively, such that cost stops being accumulated when  $G$  is reached.

By a straightforward induction on  $n$ , it is easy to show that

$$\begin{aligned} \Pr_{\mathcal{M}(\mathbf{d})}[R_{s_0 \dots s_n}] &= \Pr_{C(\mathbf{d})}[R_{s_0 \dots s_n}] && \text{for all sequences of states } s_0 \dots s_n, \text{ and} \\ \mathbb{E}_{\mathcal{M}(\mathbf{d})}^{s_{\text{in}}}[W_n] &= \mathbb{E}_{C(\mathbf{d})}^{s_{\text{in}}}[W'_n]. \end{aligned}$$

This implies that, denoting by  $N$  and  $N'$  the set of all runs of  $C$  and  $\mathcal{M}$  that do not reach the target,

$$\Pr_{C(\mathbf{d})}[N] > 0 \quad \Longleftrightarrow \quad \Pr_{\mathcal{M}(\mathbf{d})}[N'] > 0.$$

Furthermore, we get the result as

$$\begin{aligned} E_{\mathcal{M}(\mathbf{d})} &= \lim_{n \rightarrow \infty} \mathbb{E}_{\mathcal{M}(\mathbf{d})}^{s_{in}} [W_n] \\ &= \lim_{n \rightarrow \infty} \mathbb{E}_{C(\mathbf{d})}^{s_{in}} [W'_n] \\ &= E_{C(\mathbf{d})}. \end{aligned}$$

□

#### C.4 Proof of Lemma 2

In the following we denote by  $\|\mathbf{x}\|_\infty$  a *uniform norm* of a vector  $\mathbf{x}$ , i.e. the maximal absolute value of a component of  $\mathbf{x}$ . This norm induces a natural matrix norm  $\|\cdot\|_\infty$ , where for a real matrix  $X$  the number  $\|X\|_\infty$  represents the maximal absolute row sum of  $X$ .

**Lemma 2.** Let  $\alpha \in [0, 1]$  and let,  $\mathbf{d}'$  be a delay function that is  $\alpha$ -bounded by another delay function  $\mathbf{d}$ . If  $\alpha \leq \frac{1}{2 \cdot \text{Bound}[\mathbf{e}, \mathbf{d}] \cdot |S'|}$ , then

$$E_{\mathcal{M}(\mathbf{d}')} \leq E_{\mathcal{M}(\mathbf{d})} + 2 \cdot \alpha \cdot \text{Bound}[\#, \mathbf{d}] \cdot (1 + \text{Bound}[\mathbf{e}, \mathbf{d}] \cdot |S'|).$$

*Proof.* Let  $\mathbf{d}, \mathbf{d}'$  be delay functions satisfying the assumptions of the lemma. If both functions incur an infinite expected cost, the lemma clearly holds. Otherwise, due to Lemma 1 both functions incur finite cost. We start the proof by expressing the values  $E_{\mathcal{M}(\mathbf{d}')}^{\mathbf{e}}$  and  $E_{\mathcal{M}(\mathbf{d})}^{\mathbf{e}}$  as solutions of certain systems of linear equations. To achieve this, we consider discrete time absorbing Markov chains  $\mathcal{A}_{\mathbf{d}}$  and  $\mathcal{A}_{\mathbf{d}'}$  obtained by fixing strategies  $\mathbf{d}$  and  $\mathbf{d}'$  in  $\mathcal{M}$ , respectively, and replacing all transitions outgoing from states in  $G$  with self-loops on these states. Note that due to both functions incurring finite cost the set  $G$  is exactly the set of absorbing states of both  $\mathcal{A}_{\mathbf{d}}$  and  $\mathcal{A}_{\mathbf{d}'}$ .

Let  $S'' = S' \setminus G$  be the set of transient states of  $\mathcal{A}_{\mathbf{d}}$  (and hence also of  $\mathcal{A}_{\mathbf{d}'}$ ) and let  $Q_{\mathbf{d}}, Q_{\mathbf{d}'} : S'' \times S'' \rightarrow [0, 1]$  be the substochastic matrices encoding transitions in transient parts of  $\mathcal{A}_{\mathbf{d}}$  and  $\mathcal{A}_{\mathbf{d}'}$ , respectively. That is,  $Q_{\mathbf{d}}$  is an  $|S''| \times |S''|$  substochastic matrix with  $Q_{\mathbf{d}}(s, t) = T(s, \mathbf{d}(s))(t)$ , for all  $s, t \in S''$ , and similarly for  $Q_{\mathbf{d}'}$ . Next, we denote by  $f_{\mathbf{d}}, f_{\mathbf{d}'} : S'' \rightarrow \mathbb{R}_{\geq 0}$  the  $|S''|$ -dimensional vectors of one-step costs incurred with strategies  $\mathbf{d}$  and  $\mathbf{d}'$  in  $\mathcal{M}$  and  $\mathcal{M}^*$ , respectively. That is,  $f_{\mathbf{d}}(s) = \mathbf{e}(s, \mathbf{d}(s))$  for all  $s \in S''$ , and similarly for  $f_{\mathbf{d}'}$ . By our assumptions, we have  $Q_{\mathbf{d}'} = Q_{\mathbf{d}} + \Delta Q$ , where  $\Delta Q$  is a matrix whose each entry has absolute value bounded by  $\alpha$ , and  $f_{\mathbf{d}'} = f_{\mathbf{d}} + \Delta f$ , where  $\Delta f$  is a vector whose each entry is  $\leq \alpha$ .

From standard results on Markov models with total accumulated reward criterion (e.g. [34, Theorem 7.1.3]) we have that the  $|S''|$ -dimensional vector  $(E_{\mathcal{M}[\mathbf{s}](\mathbf{d})}^{\mathbf{e}})_{s \in S''}$  is equal to the unique solution  $\mathbf{x}$  of the following system of linear equations:

$$(I - Q_{\mathbf{d}}) \cdot \mathbf{x} = f_{\mathbf{d}}.$$

(The uniqueness of the solution comes from  $\mathcal{A}_{\mathbf{d}}$  being absorbing, which means that  $(I - Q_{\mathbf{d}})$  is invertible [24]). Similarly, the vector  $(E_{\mathcal{M}[\mathbf{s}](\mathbf{d}')}^{\mathbf{e}})_{s \in S''}$  is the unique solution  $\mathbf{y}$  of the following system of equations:

$$(I - Q_{\mathbf{d}} - \Delta Q) \cdot \mathbf{y} = f_{\mathbf{d}} + \Delta f.$$

Hence, it suffices to obtain a bound on the maximal positive component of  $\mathbf{x} - \mathbf{y}$ . We first replace all negative components of  $\Delta f$  with zeroes and thus obtain a new vector  $\Delta f'$ . We claim that the unique solution  $\mathbf{y}'$  to the system

$$(I - Q_{\mathbf{d}} - \Delta Q) \cdot \mathbf{y}' = f_{\mathbf{d}} + \Delta f'.$$

satisfies  $\mathbf{y}' \geq \mathbf{y}$ . Indeed, this is because  $\mathbf{y}' = (I - Q_{\mathbf{d}} - \Delta Q)^{-1} \cdot (f_{\mathbf{d}} + \Delta f') \geq (I - Q_{\mathbf{d}} - \Delta Q)^{-1} \cdot (f_{\mathbf{d}} + \Delta f) = \mathbf{y}$ , the middle inequality following that the matrix  $(I - Q_{\mathbf{d}} - \Delta Q)^{-1}$ , so called fundamental matrix of a Markov chain  $\mathcal{A}_{\mathbf{d}}$ , is always non-negative. So to prove the lemma it suffices to give a bound on  $\|\mathbf{x} - \mathbf{y}'\|_{\infty}$ . Note that  $\|\Delta f'\|_{\infty} \leq \alpha$ .

We use standard results on stability of perturbed systems of linear equations. From, e.g. proof of Theorem 3 in [20] it follows that

$$\|\mathbf{x} - \mathbf{y}'\|_{\infty} \leq \frac{\|(I - Q_{\mathbf{d}})^{-1}\|_{\infty}}{1 - \|(I - Q_{\mathbf{d}})^{-1}\|_{\infty} \cdot \|\Delta Q\|_{\infty}} \cdot (\|\Delta f'\|_{\infty} + \|\mathbf{x}\|_{\infty} \cdot \|\Delta Q\|_{\infty}). \quad (2)$$

We know that  $\|\Delta Q\|_{\infty} \leq \alpha \cdot |S'|$  and  $\|\Delta f'\|_{\infty} \leq \alpha$ , and from the definition of  $\mathbf{x}$  we have that  $\|\mathbf{x}\|_{\infty} = \text{Bound}[\epsilon, \mathbf{d}]$ . Moreover, as noted above,  $(I - Q_{\mathbf{d}})^{-1}$  is the fundamental matrix of the absorbing Markov chain  $\mathcal{A}_{\mathbf{d}}$ , i.e. a matrix such that the entry  $(I - Q_{\mathbf{d}})^{-1}(s, t)$  is equal to the expected number of visits to a transient state  $t$  when starting in a transient state  $s$  of  $\mathcal{A}_{\mathbf{d}}$ . In particular, sum of each row of  $(I - Q_{\mathbf{d}})^{-1}$  is bounded by  $\text{Bound}[\#, \mathbf{d}]$ , and thus  $\|(I - Q_{\mathbf{d}})^{-1}\|_{\infty} \leq \text{Bound}[\#, \mathbf{d}]$ . Plugging these (in)equalities into (2) we get

$$\begin{aligned} \|\mathbf{x} - \mathbf{y}'\|_{\infty} &\leq \frac{\text{Bound}[\#, \mathbf{d}]}{1 - \text{Bound}[\#, \mathbf{d}] \cdot \alpha \cdot |S'|} \cdot (\alpha + \text{Bound}[\epsilon, \mathbf{d}] \cdot \alpha \cdot |S'|) \\ &\leq 2 \cdot \text{Bound}[\#, \mathbf{d}] \cdot (\alpha + \text{Bound}[\epsilon, \mathbf{d}] \cdot \alpha \cdot |S'|), \end{aligned}$$

the last inequality following from our assumption that  $\alpha \leq 1/(2 \cdot \text{Bound}[\#, \mathbf{d}] \cdot |S'|)$ . The lemma easily follows.  $\square$

### C.5 Proof of Lemma 3

**Lemma 3.** There are positive numbers  $D_1, D_2 \in \exp(\|C\|^{O(1)})$  computable in time polynomial in  $\|C\|$  such that the following holds for any  $\alpha \in [0, 1]$  and any delay function  $\mathbf{d}$ : If we put

$$\delta := \alpha/D_1 \quad \text{and} \quad \bar{d} := |\log(\alpha)| \cdot D_2 \cdot \text{Bound}[\epsilon, \mathbf{d}],$$

then  $D(\delta, \bar{d})$  contains a delay function which is  $\alpha$ -bounded by  $\mathbf{d}$ .

The proof of the lemma proceeds in 2 phases. First we show how to set  $\delta$  in such a way that for every delay function  $\mathbf{d}$  there is a function using only multiples of  $\delta$  which is  $\alpha/2$ -bounded by  $\mathbf{d}$ . In the second phase we show how to choose  $\bar{d}$  such that for each delay function  $\mathbf{d}$  (and in particular each delay function that uses only multiples of  $\delta$ ) there is a function  $\alpha/2$ -bounded by  $\mathbf{d}$  that does not use action greater than  $\bar{d}$ .

We will need the following definition.



**Definition 7.** For a fdCTMC structure  $(S, \lambda, P, S_{fd}, F, s_{in})$ , state  $s \in S'$ , set of states  $S' \subseteq S$  and delay  $d > 0$  we define a fdCTMC  $C'[s](d)$ , where  $C' = (S \cup \{\bar{s}\}, \lambda, P', S_{fd}, F', \bar{s})$ ,  $\bar{s}$  is new initial state, that was before not in  $S$ , and for all  $s, s' \in S$  holds

$$P'(s', s'') = \begin{cases} P(s, s'') & \text{if } s' = \bar{s}, \\ P(s', s'') & \text{if } s' \notin S', \\ 1 & \text{if } s' = s'' \in S', \\ 0 & \text{otherwise,} \end{cases}$$

and

$$F'(s', s'') = \begin{cases} F(s, s'') & \text{if } s' = \bar{s}, \\ F(s', s'') & \text{if } s' \notin S', \\ 1 & \text{if } s' = s'' \in S', \\ 0 & \text{otherwise.} \end{cases}$$

**Lemma 7.** There is a number  $D_1 \in \exp(\|C\|^{O(1)})$  computable in time polynomial in  $\|C\|$  such that for every  $\alpha$  and every delay function  $\mathbf{d}$  the following holds: If we put  $\delta = \alpha/D_1$ , then  $\{\mathbf{d} \mid \forall s \in S' \exists k \in \mathbb{N} : \mathbf{d}(s) = k\delta\}$  contains a delay function which is  $(\alpha/2)$ -bounded by  $\mathbf{d}$ .

*Proof.* Let us fix arbitrary strategy  $\mathbf{d}$  of  $\mathcal{M}$ . We show that for each state  $s \in S'$  there is strategy  $\mathbf{d}^* \in \{\mathbf{d} \mid \forall s \in S' \exists k \in \mathbb{N} : \mathbf{d}(s) = k\delta\}$  such that for all  $s, s' \in S'$  it holds that

$$|T(s, \mathbf{d}(s))(s') - T(s, \mathbf{d}^*(s))(s')| \leq \lambda \cdot \delta.$$

Please observe that to compute  $T(s, d)(s')$  we can perform the transient analysis of  $C'[s](d)$  in time  $d$  and then perform in zero time fixed delay transitions. The  $C'[s](d)$  performs as continuous time Markov chain (CTMC) until time  $d$ , thus we can use methods to compute the transient analysis of CTMC in time  $d$ . One of them is to use matrix exponential  $P''(t) = e^{tQ}$  [32] for a  $Q$ -matrix for  $C'[s](\mathbf{d}(s))$ . It is known that  $P''(t)_{s,s'}$  is the probability of moving from state  $s$  to  $s'$  exactly in time  $t$ . Moreover for  $t \geq 0$  it holds that

$$\frac{d P''(t)}{dt} = P''(t) \cdot Q.$$

Since the sum of positive entries in each line in  $Q$  is at most the uniformization rate  $\lambda$  and that  $P''(t)$  for every  $t \geq 0$  is correct transition kernel it holds that for all  $s, s' \in S'$  and  $t \geq 0$

$$(P''(t) \cdot Q)_{s,s'} \leq \lambda.$$

Thus the transient probability at time  $t > 0$  in  $C'[s](\mathbf{d}(s))$  can change at most  $\lambda \cdot \delta$  if we move time  $\mathbf{d}(s)$  by  $\delta$ . Since  $T$  is also transition kernel it does only convex combination of transient probability at time  $\mathbf{d}(s) \pm \delta$  thus it cannot increase the error. Altogether we have that for all  $s, s' \in S'$

$$|T(s, \mathbf{d}(s))(s') - T(s, \mathbf{d}^*(s))(s')| \leq \lambda \cdot \delta. \quad (3)$$

Hence, the strategy  $\mathbf{d}^*$  we seek can be obtained by simply rounding the components  $\mathbf{d}$  to the nearest positive integer multiple of  $\delta$ .

Now we show that for each state  $s \in S'$  it holds

$$|\mathbb{E}(s, \mathbf{d}(s)) - \mathbb{E}(s, \mathbf{d}^*(s))| \leq 2 \cdot \lambda \cdot \delta \cdot \max R + \delta \cdot \max R \leq 2 \cdot (\lambda + 1) \cdot \delta \cdot \max R.$$

Recall that  $\mathbb{E}(s, d) = E_{C[s](d)}^{Cost[S']}$ , i.e.  $\mathbb{E}(s, d)$  is the expected cost accumulated until hitting state from  $S'$ . This cost can be divided into expected cost accumulated by rate costs and by expected cost accumulated by impulse costs. First we bound the change on expected rate cost until hitting  $S'$ . Let  $O$  be random variable assigning run the time until hitting state from  $S'$ . Assume that the delay  $\mathbf{d}(s)$  was increased by  $\delta$  (if it was decreased the argument is symmetric- we only start with  $\mathbf{d}^*(s)$ ). Then the time until hitting  $S'$  by each run is increased at most by  $\delta$ . Thus the maximal change in expected time until hitting  $S'$  is  $\delta$ . The maximum rate cost in every state is bounded by  $\max R$ . Thus the maximal change in expected rate cost accumulated until time  $\mathbf{d}(s)$  if we increase it by  $\delta$  is  $\delta \cdot \max R$ .

Similarly, if the delay of  $\mathbf{d}(s)$  is increased by  $\delta$  (if it was decreased the argument is symmetric- we just start with  $\mathbf{d}^*(s)$ ) we give every run a chance to do some more steps, what can increase the impulse cost. The increase in expected impulse cost until hitting state from  $S'$  is bounded by the change in expected number of exponential steps from  $s''$  in time  $\mathbf{d}(s)$  plus change in expected number of fixed-delay steps in time  $\mathbf{d}(s)$  multiplied by the maximum impulse cost. The expected number of exponential steps with rate  $\lambda$  in time  $\mathbf{d}(s)$  is mean of Poisson distribution with parameter  $\lambda \cdot \mathbf{d}(s)$ , what equals  $\lambda \cdot \mathbf{d}(s)$ . Thus if the  $\mathbf{d}(s)$  was increased by  $\delta$  the increase in expected number exponential steps is  $\lambda \cdot \delta$ . Please observe, that only one fixed-delay step can happen until state from  $S'$  is reached, because all fixed-delay transitions go to states in  $S'$ . The fixed-delay transition fires in time  $\mathbf{d}(s)$  if the run is still in state from  $S_{fd}$ . If the time is increased by  $\delta$  the probability of run to be in states  $S_{fd}$  in time  $\mathbf{d}(s)$  can change at most by  $\lambda \cdot \delta$  (from bounded change in transient probability proved in first paragraph of this proof). Thus the maximal increase in expected impulse cost until reaching  $S'$  from  $s$  is  $\max R \cdot (\lambda \cdot \delta + \lambda \cdot \delta)$ . Altogether we have that

$$|\mathbb{E}(s, \mathbf{d}(s)) - \mathbb{E}(s, \mathbf{d}^*(s))| \leq 2 \cdot \lambda \cdot \delta \cdot \max R + \delta \cdot \max R \leq 2 \cdot (\lambda + 1) \cdot \delta \cdot \max R. \quad (4)$$

From (3) and (4) it follows that to get the lemma it suffices to put  $D_1 = \max\{\lambda, 2(\lambda + 1) \cdot \max R\}$ .

**Lemma 8.** *There is a number  $D_2 \in \exp(\|C\|^{O(1)})$  computable in time polynomial in  $\|C\|$  such that for every  $\alpha$  and every delay function  $\mathbf{d}$  the following holds: If we put  $\bar{d} := \lceil \log(\alpha) \rceil \cdot D_2 \cdot \text{Bound}[\mathbb{E}, \mathbf{d}]$ , then  $\{\mathbf{d} \mid \forall s \in S' \mathbf{d}(s) \leq \bar{d}\}$  contains a delay function which is  $(\alpha/2)$ -bounded by  $\mathbf{d}$ .*

*Proof.* Let us fix arbitrary strategy  $\mathbf{d}$  of  $\mathcal{M}$  and  $s \in S'$ . We use the  $C'[s](\mathbf{d}(s))$  from Definition 7 (that can be used to compute  $T(s, \mathbf{d}(s))$  using transient analysis of continuous time Markov chains in time  $\mathbf{d}(s)$ ) to generate a directed graph  $Gr(S \cup \{\bar{s}\}, tra)$ , where  $(s, s') \in tra$  iff  $P'(s, s') > 0$ . Let  $st \subseteq S \cup \{\bar{s}\}$  be set of states that are reachable in  $Gr$  from  $\bar{s}$ . We define directed graph  $gr = (st, tra')$  to be induced graph by  $st$  from  $Gr$ . We distinguish two cases:

- $C'[s](\mathbf{d}(s))$  contains bottom strongly connected component (BSCC)  $O \subseteq S \cup \{\bar{s}\}$  in  $Gr$ , such that  $O \cap S' = \emptyset$ , and
- all BSCCs in  $Gr$  are self-loops and contain only states from  $S'$ .

Observe that other cases cannot happen from definition of  $C'[s](\mathbf{d}(s))$ . First we take care of first item. We show that it must hold

$$\mathbf{d}(s) \leq \frac{\text{Bound}[\epsilon, \mathbf{d}]}{\min P^{|\mathcal{S}_{\text{fd}}|} \cdot \min R}, \quad (5)$$

which gives us a simple bound on  $\bar{d}$ . The minimal probability to reach  $O$  is the  $\min P^{|\mathcal{S}_{\text{fd}}|}$  because the minimal path to  $O$  in  $Gr$  can have length at most  $|\mathcal{S}_{\text{fd}}|$  and the minimal positive branching probability is  $\min P$ . Observe that if  $O$  is reached by a run it will stay there until the fixed-delay transition is fired. The expected rate cost accumulated by paths that reach  $O$  is at least  $\mathbf{d}(s) \cdot \min R$ . Hence we gradually get

$$\begin{aligned} \min P^{|\mathcal{S}_{\text{fd}}|} \cdot \min R \cdot \mathbf{d}(s) &\leq \text{Bound}[\epsilon, \mathbf{d}] \\ \mathbf{d}(s) &\leq \frac{\text{Bound}[\epsilon, \mathbf{d}]}{\min P^{|\mathcal{S}_{\text{fd}}|} \cdot \min R}, \end{aligned}$$

as required.

From now we assume that  $gr$  of  $C'[s](\mathbf{d}(s))$  for state  $s$  does contain only BSCCs reachable from  $s$  that are self-loops and all contain only states in  $S'$ . We need to obtain a suitable bound on  $\bar{d}$  for this case. Obviously by restricting strategies that do not have large delays we can only decrease the expected one step cost. So it suffices to construct  $\bar{d}$  such that for each  $\mathbf{d}$  there is a strategy  $\mathbf{d}^* \in \{\mathbf{d} \mid \forall s \in S' \mathbf{d}(s) \leq \bar{d}\}$  such that for all  $s' \in S'$  it holds that

$$|T(s, \mathbf{d}(s))(s') - T(s, \mathbf{d}^*(s))(s')| \leq \alpha/2.$$

Obviously it holds that  $|T(s, \mathbf{d}(s))(s') - T(s, \mathbf{d}^*(s))(s')| \leq |1 - \sum_{s' \in S'} T(s, \mathbf{d}^*(s))(s')|$ . Observe that  $|1 - \sum_{s' \in S'} T(s, \mathbf{d}^*(s))(s')|$  denotes the probability that a state from  $S'$  was not reached yet, i.e. that run in  $C'[s](\mathbf{d}(s))$  is still in transient part of  $C'[s](\mathbf{d}(s))$ . Obviously as  $\mathbf{d}^*(s)$  approaches infinity  $|1 - \sum_{s' \in S'} T(s, \mathbf{d}^*(s))(s')|$  goes to zero. We will overestimate  $|1 - \sum_{s' \in S'} T(s, \mathbf{d}^*(s))(s')|$  by overestimating the probability that any BSCC has not yet been reached until time  $t > 0$  in  $C'[s](\mathbf{d}(s))$ . We divide  $t$  into steps of length  $|\mathcal{S}_{\text{fd}}|/\lambda$ . Using uniformization method we get:

$$\begin{aligned} \left| 1 - \sum_{s' \in S'} T(s, \frac{|\mathcal{S}_{\text{fd}}|}{\lambda})(s') \right| &= \sum_{s' \in S'} e^{-\lambda \frac{|\mathcal{S}_{\text{fd}}|}{\lambda}} \sum_{i=0}^{\infty} P^i(s, s') \cdot \frac{(\lambda \frac{|\mathcal{S}_{\text{fd}}|}{\lambda})^i}{i!} \\ &= \sum_{s' \in S'} e^{-|\mathcal{S}_{\text{fd}}|} \sum_{i=0}^{\infty} P^i(s, s') \cdot \frac{|\mathcal{S}_{\text{fd}}|^i}{i!} \\ &\geq \sum_{s' \in S'} e^{-|\mathcal{S}_{\text{fd}}|} P^{|\mathcal{S}_{\text{fd}}|}(s, s') \cdot \frac{|\mathcal{S}_{\text{fd}}|^{|\mathcal{S}_{\text{fd}}|}}{|\mathcal{S}_{\text{fd}}|!} \\ &\geq \sum_{s' \in S'} e^{-|\mathcal{S}_{\text{fd}}|} P^{|\mathcal{S}_{\text{fd}}|}(s, s') \cdot \frac{|\mathcal{S}_{\text{fd}}|^{|\mathcal{S}_{\text{fd}}|}}{|\mathcal{S}_{\text{fd}}|^{|\mathcal{S}_{\text{fd}}|}} \\ &\geq e^{-|\mathcal{S}_{\text{fd}}|} \min P^{|\mathcal{S}_{\text{fd}}|} = \left( \frac{\min P}{e} \right)^{|\mathcal{S}_{\text{fd}}|}. \end{aligned}$$

The probability that in time  $\mathbf{d}^\star(s) = l \cdot |S_{\text{fid}}|/\lambda$  that state from  $S'$  has still not been reached is

$$\begin{aligned} \left| 1 - \sum_{s' \in S'} T(s, \mathbf{d}^\star(s))(s') \right| &\leq \left( 1 - \left( \frac{\min P}{e} \right)^{|S_{\text{fid}}|} \right)^l \\ &= \left( 1 - \left( \frac{\min P}{e} \right)^{|S_{\text{fid}}|} \right)^{\frac{\mathbf{d}^\star(s) \cdot \lambda}{|S_{\text{fid}}|}} \end{aligned}$$

To ensure that the right-hand side of the above inequality is  $\leq \alpha/2$ , it suffices to ensure that

$$\mathbf{d}^\star(s) \geq \frac{|\log(\alpha/2)| \cdot |S| \cdot e^{|S|}}{\lambda \cdot \min P^{|S|}}. \quad (6)$$

Combining (5) and (6) we get that it suffices to put  $D_2 = (|S| \cdot e^{|S|})/(\min P^{|S|} \cdot \min\{1, \lambda\} \cdot \min\{1, \min R\})$ .

## C.6 Proof of Lemma 4

**Lemma 4.** There is a positive number  $N \in \exp(\|C\|^{O(1)})$  computable in time polynomial in  $\|C\|$  such that the following holds: for any  $\varepsilon' > 0$ , there is a globally  $\varepsilon'/2$ -optimal delay function  $\mathbf{d}'$  with

$$\text{Bound}[\#, \mathbf{d}'] \leq \frac{\text{Bound}[\epsilon, \mathbf{d}']}{\varepsilon'} \cdot N. \quad (7)$$

We start the proof with a lemma which shows that for every  $\varepsilon > 0$  there is a *globally*  $\varepsilon$ -optimal delay function in  $\mathcal{M}$ .

**Lemma 9.** For every  $\varepsilon > 0$  there is a globally  $\varepsilon$ -optimal delay function in  $\mathcal{M}$ .

*Proof.* Fix  $\varepsilon > 0$ . For every state  $s$  there is trivially a delay function  $\mathbf{d}_s$  which is  $(\varepsilon/2)$ -optimal in  $s$ . Fix such an  $(\varepsilon/2)$ -optimal function for each state  $s$ . Since there are only finitely many  $s$ , there is a number  $\alpha > 0$  such that

$$2 \cdot \alpha \cdot \text{Bound}[\#, \mathbf{d}_s] \cdot (1 + \text{Bound}[\epsilon, \mathbf{d}_s] \cdot |S'|) \leq \varepsilon/2 \quad (8)$$

for all  $s$ . Moreover, from Lemma 3 it follows that for each  $s$  there is sufficiently small  $\delta_s$  and sufficiently large  $\bar{d}_s$  such that  $D(\delta_s, \bar{d}_s)$  contains a function that is  $\alpha$ -bounded by  $\mathbf{d}_s$ . Let  $\delta = \min_{s \in S_{\text{fid}}} \delta_s$  and  $\bar{d} = \max_{s \in S_{\text{fid}}} \bar{d}_s$ . Then for each  $s$ , the set  $D(\delta, \bar{d})$  contains a delay function that is  $\alpha$ -bounded by  $\mathbf{d}_s$ . From (8) and Lemma 2 it follows that in the *finite* DTMDP obtained from  $\mathcal{M}$  by restricting the set of actions to  $\{k\delta \mid k \in \mathbb{N}, k > 0, k\delta \leq \bar{d}\} \cup \{\infty\}$  the infimum expected cost achievable from a given state is at most  $\varepsilon/2$  away from the infimum cost achievable from this state in  $\mathcal{M}$ . Moreover, in a finite DTMDP with expected total accumulated cost objective there is always a memoryless delay function<sup>5</sup> that is optimal in every vertex [34]. Hence,  $D(\delta, \bar{d})$  contains a delay function that is globally  $\varepsilon$ -optimal in  $\mathcal{M}$ .  $\square$

<sup>5</sup> I.e. a delay function whose decision is based only on the current vertex, as defined in Section 2.

Recall that  $S' =$  is the set of states of DTMDP  $\mathcal{M}$ , i.e. the set of those states of fdCTMC structure  $C$  in which the fixed-delay is newly set or switched off, together with all goal states from  $G$ . An *fd-skeleton* of  $\mathcal{M}$  is the directed graph  $(S', E)$  such that  $(s, t) \in E$  if and only if  $s \notin G$  and  $F(s, t) > 0$ . A *sink* of the DTMDP  $\mathcal{M}$  is a bottom strongly connected component of its fd-skeleton, i.e. a set of states  $K \subseteq S'$  such that a) the subgraph of  $(S', E)$  induced by  $K$  is strongly connected, and b) whenever  $(s, t) \in E$  and  $s \in K$ , then also  $t \in K$ .

Let  $\mathbf{d}$  be any delay function in  $\mathcal{M}$ . We say that a sink  $K$  is *bad* for  $\mathbf{d}$  if the following two conditions hold:

- $I_F(s, t) = 0$  for all  $s, t \in K$ , and
- $\mathbf{d}(s) < \underline{d}$  for all  $s \in K$ , where  $\underline{d}$  is a suitable number defined below.

A sink is *good* for  $\mathbf{d}$  if it is not bad for  $\mathbf{d}$ .

Let  $\underline{d}$  be a number such that it holds

$$e^{\lambda \underline{d}} = 1 + \frac{\varepsilon \cdot \min R}{4 \cdot \lambda \cdot (\overline{\text{Val}}[\mathcal{M}] + \varepsilon + \max R \cdot (2 + 2/\lambda))^2}. \quad (9)$$

Note that using Taylor series for  $\ln(1 + z) = z - z^2/2 + z^3/3 - \dots$  we get

$$\underline{d} \leq \frac{\varepsilon \cdot \min R}{4 \cdot \lambda^2 \cdot (\overline{\text{Val}}[\mathcal{M}] + \varepsilon + \max R \cdot (2 + 2/\lambda))^2}.$$

**Lemma 10.** *Let  $\varepsilon > 0$  be arbitrary, and let  $\mathbf{d}$  be any globally  $\varepsilon$ -optimal delay function in  $\mathcal{M}$  such that all sinks of  $\mathcal{M}$  are good for  $\mathbf{d}$ . Then for all pairs of states  $s, t$  of  $\mathcal{M}$  it holds*

$$\mathbb{E}_{\mathcal{M}[s](\mathbf{d})}[\#_t^G] \leq \frac{|S_{\text{fd}}|}{(\min P \cdot e^{-\lambda \underline{d}})^{|S_{\text{fd}}|}} \cdot \frac{\overline{\text{Val}}[\mathcal{M}] + \varepsilon}{\min P \cdot \min R \cdot (1 - e^{-\lambda \underline{d}})/\lambda}. \quad (10)$$

*Proof.* First we define *good states* and show that there is lower bound on cost paid in one step from a good state. Because we have upper bound on cost from any state  $\overline{\text{Val}}[\mathcal{M}] + \varepsilon$  we cannot visit too many times good states before reaching goal state. We show that we have small expected number of steps from any state to reach a good state. Finally we put all the previous results together to show the final result.

We define *good states*  $L \subseteq S' \setminus G$  for  $\mathcal{M}$  and  $\mathbf{d}$  as  $L = (S' \setminus (S_{\text{fd}} \cup G)) \cup \{s \in S' \setminus G \mid \mathbf{d}(s) \geq \underline{d} \text{ or exists } t \in S' \text{ such that } I_F(s, t) > 0 \text{ and } F(s, t) > 0\}$ . Now we show that there is lower bound on one step cost from good states, i.e. for all  $s \in L$  holds  $\mathbb{E}(s, \mathbf{d}(s)) \geq a$  for some  $a > 0$ . In states  $S' \setminus S_{\text{fd}}$  the minimal expected cost accumulated in one step is  $\min R/\lambda$ , since the expected time to execute exponential transition is  $1/\lambda$  and minimal rate cost is  $\min R$ . The expected cost accumulated in one step from state  $s \in \{s \in S' \setminus G \mid \mathbf{d}(s) \geq \underline{d}\}$  can

be bounded as follows:

$$\begin{aligned}
\mathbb{E}(s, \mathbf{d}(s)) &\geq \min R \cdot (\text{expected time spent in } s \text{ until the first transition is taken}) \\
&= \min R \cdot \int_0^{\mathbf{d}(s)} t \cdot \lambda \cdot e^{-\lambda t} dt + \mathbf{d}(s) \cdot e^{-\lambda \mathbf{d}(s)} \\
&= \min R \cdot \left( -\mathbf{d}(s) \cdot e^{-\lambda \mathbf{d}(s)} + \frac{1 - e^{-\lambda \mathbf{d}(s)}}{\lambda} + \mathbf{d}(s') \cdot e^{-\lambda \mathbf{d}(s)} \right) \\
&= \min R \cdot \frac{1 - e^{-\lambda \mathbf{d}(s)}}{\lambda} \\
&\geq \min R \cdot \frac{1 - e^{-\lambda \underline{d}}}{\lambda} \\
&\geq \min R \cdot \min P \cdot \frac{1 - e^{-\lambda \underline{d}}}{\lambda},
\end{aligned} \tag{11}$$

where the inequality (11) follows from the fact that  $s \in \{s \in S' \setminus G \mid \mathbf{d}(s) \geq \underline{d}\}$ . Finally if  $s \in \{s \in S' \setminus G \mid \mathbf{d}(s) < \underline{d}\}$  and exists  $t \in S'$  such that  $\mathcal{I}_F(s, t) > 0$  and  $F(s, t) > 0$  then it holds that

$$\begin{aligned}
\mathbb{E}(s, \mathbf{d}(s)) &\geq \min R \cdot \min P \cdot e^{-\lambda \mathbf{d}(s)} \\
&\geq \min R \cdot \min P \cdot e^{-\lambda \underline{d}},
\end{aligned}$$

where the first inequality follows from fact that probability of taking the transition with nonzero impulse cost (thus at least  $\min R$ ) is at least minimal branching probability  $\min P$  times the probability that fixed delay transition is taken what is  $e^{-\lambda \mathbf{d}(s)}$ . The last inequality follows from the fact that  $s \in \{s \in S' \setminus G \mid \mathbf{d}(s) < \underline{d}\}$  and exists  $t \in S'$  such that  $\mathcal{I}_F(s, t) > 0$  and  $F(s, t) > 0$ . Obviously for every  $s \in L$  holds that

$$\begin{aligned}
\mathbb{E}(s, \mathbf{d}(s)) &\geq \min \left\{ \min R \cdot \min P \cdot e^{-\lambda \underline{d}}, \min R / \lambda, \min R \cdot \min P \cdot \frac{1 - e^{-\lambda \underline{d}}}{\lambda} \right\} \\
&= \min R \cdot \min P \cdot \frac{1 - e^{-\lambda \underline{d}}}{\lambda}.
\end{aligned}$$

Now provide an upper bound  $\overline{W}$  on the expected number of steps to reach a good or goal state (i.e.  $L \cup G$ ) from any  $s \in S'$ :

$$\begin{aligned}
\max_{s \in S'} \mathbb{E}_{\mathcal{M}(s)(\mathbf{d})} [\#_{S'}^{L \cup G}] &\leq \\
&= \sum_{i=1}^{\infty} i \cdot |S_{\text{fd}}| \cdot (1 - \min_{s \in S'} \Pr_{\mathcal{M}(\mathbf{d})}(\text{reach } L \cup G \text{ from } s \text{ in } |S_{\text{fd}}| \text{ steps}))^{i-1} \\
&\quad \cdot \min_{s \in S'} \Pr_{\mathcal{M}(\mathbf{d})}(\text{reach } L \cup G \text{ from } s \text{ in } |S_{\text{fd}}| \text{ steps}) \\
&\leq \sum_{i=1}^{\infty} i \cdot |S_{\text{fd}}| \cdot (1 - (\min P \cdot e^{-\lambda \underline{d}})^{|S_{\text{fd}}|})^{i-1} \cdot (\min P \cdot e^{-\lambda \underline{d}})^{|S_{\text{fd}}|} \\
&= \frac{|S_{\text{fd}}|}{(\min P \cdot e^{-\lambda \underline{d}})^{|S_{\text{fd}}|}} = \overline{W},
\end{aligned} \tag{12}$$

$$\tag{13}$$

where (13) follows from the fact that we can do at most  $|S_{\text{fd}}|$  steps avoiding to reach some of the  $L \cup G$  with positive probability, by moving through states  $S_{\text{fd}}$  (states not belonging to  $S_{\text{fd}}$  are good states from definition of good states). This probability is at least the minimal branching probability  $\text{minP}$  times probability  $e^{-\lambda d(s)}$  that fixed delay transition was fired for  $s \in S_{\text{fd}} \setminus L$ , what is at least  $e^{-\lambda d}$ , because  $d(s) < d$  from definition of good states.

To get the final result we have to define few random variables. For a DTMDP  $\mathcal{M}$  and set  $L \subseteq S'$  we define the series of random variables  $(W_n^L)$ , where for each  $n \in \mathbb{N}$  the  $W_n^L : (S' \cdot \text{Act})^\omega \rightarrow \mathbb{N}_0$  gives the number of steps between  $n$ th and  $(n+1)$ st visit of  $L$  before reaching  $G$ , i.e.

$$W_n^L(s_0 a_1 s_1 a_2 s_2 \dots) = \begin{cases} H_n^L & \text{if } n = 1, \\ H_n^L - H_{n-1}^L & \text{if } n > 1. \end{cases}$$

Obviously it holds that

$$\#_{S'}^G = \sum_{n=1}^{\infty} W_n^L$$

and for each  $n \in \mathbb{N}$  and  $s \in S'$  it holds that

$$\mathbb{E}_{\mathcal{M}[s](\mathbf{d})}[W_{n+1}^L] = \sum_{s' \in L} \Pr_{\mathcal{M}(\mathbf{d})}(X_{H_n^L}^{\mathcal{M}} = s') \cdot \mathbb{E}_{\mathcal{M}[s](\mathbf{d})}[W_{n+1}^L \mid X_{H_n^L}^{\mathcal{M}} = s'] \leq \overline{W} \cdot \Pr_{\mathcal{M}(\mathbf{d})}(X_{H_n^L}^{\mathcal{M}} \in L),$$

where  $\Pr_{\mathcal{M}(\mathbf{d})}(X_{H_n^L}^{\mathcal{M}} \in L)$  is probability of at least  $n$  times reaching a state from  $L$  before reaching any of the goal states  $G$ . We define a random variable  $V$  denoting number of visits of  $L$  before hitting  $G$ , i.e.

$$V = \sum_{i=0}^{H_1^0} 1_{X_i \in L},$$

where 1 is indicator function. Obviously it holds that

$$\mathbb{E}_{\mathcal{M}[s](\mathbf{d})}[W_{n+1}^L] \leq \overline{W} \cdot \Pr_{\mathcal{M}(\mathbf{d})}(X_{H_n^L}^{\mathcal{M}} \in L) = \overline{W} \cdot \Pr_{\mathcal{M}(\mathbf{d})}^s(V \geq n).$$

Also it holds that

$$\mathbb{E}_{\mathcal{M}[s](\mathbf{d})}[W_1^L] = \mathbb{E}_{\mathcal{M}[s](\mathbf{d})}[\#_{S'}^{L \cup G}] \leq \overline{W} = \overline{W} \cdot \Pr_{\mathcal{M}(\mathbf{d})}^s(V \geq 0).$$

Altogether we have

$$\begin{aligned} \mathbb{E}_{\mathcal{M}[s](\mathbf{d})}[\#_{S'}^G] &= \sum_{n=1}^{\infty} \mathbb{E}_{\mathcal{M}[s](\mathbf{d})}[W_n^L] \\ &\leq \overline{W} \cdot \Pr_{\mathcal{M}(\mathbf{d})}^s(V \geq n) \\ &= \overline{W} \cdot \mathbb{E}_{\mathcal{M}[s](\mathbf{d})}[V] \\ &\leq \overline{W} \cdot \frac{\overline{\text{Val}}[\mathcal{M}] + \varepsilon}{\text{minP} \cdot \text{minR} \cdot (1 - e^{-\lambda d})/\lambda} \\ &\leq \frac{|S_{\text{fd}}|}{(\text{minP} \cdot e^{-\lambda d})^{|S_{\text{fd}}|}} \cdot \frac{\overline{\text{Val}}[\mathcal{M}] + \varepsilon}{\text{minP} \cdot \text{minR} \cdot (1 - e^{-\lambda d})/\lambda}, \end{aligned} \tag{14}$$

where (14) follows from the fact that states  $L$  cannot be visited too often because in each visit at least  $\min P \cdot \min R \cdot (1 - e^{-\lambda d})/\lambda$  of cost is accumulated and the overall bound on cost is  $\overline{\text{Val}}[\mathcal{M}] + \varepsilon$ . Finally trivially for all  $s, t \in S'$  holds  $\mathbb{E}_{\mathcal{M}[s](\mathbf{d})}[\#_t^G] \leq \mathbb{E}_{\mathcal{M}[s](\mathbf{d})}[\#_{S'}^G]$ .

The previous lemma shows that to prove Lemma 4 it suffices to show that there is an  $\varepsilon$ -optimal delay function for which all sinks of  $\mathcal{M}$  are good. Indeed, plugging the definition of  $\underline{d}$  into (10) we get that such a delay function  $\mathbf{d}$  satisfies

$$\begin{aligned} \text{Bound}[\#, \mathbf{d}] &\leq \frac{\text{Bound}[\mathbb{E}, \mathbf{d}]}{\varepsilon} \cdot \frac{(\overline{\text{Val}}[\mathcal{M}] + \max R \cdot (2 + 2/\lambda) \cdot \lambda)^2 \cdot 4\lambda^2 \cdot e^{2\lambda d}}{\min R^2 \cdot \min P} \\ &\leq \frac{\text{Bound}[\mathbb{E}, \mathbf{d}]}{\varepsilon} \cdot \frac{(\overline{\text{Val}}[\mathcal{M}] + \max R \cdot (2 + 2/\lambda) \cdot \lambda)^2 \cdot 4\lambda^2 \cdot 4}{\min R^2 \cdot \min P}, \end{aligned} \quad (15)$$

the last inequality following from the fact that  $e^{\lambda d} \leq 2$  (the numerator in (9) is at most  $\min R$  while the denominator is at least  $\max R \cdot (4/\lambda) \cdot 4\lambda \geq \max R \geq \min R$ ). The number  $N$  from Lemma 4 can be easily obtained from the right-hand side of (15) and from Lemma 5.

The crucial idea, which is summarized in the following lemma, is that any  $\varepsilon$ -optimal delay vector can be modified (by “inflating” one of its components to  $\underline{d}$ ) in such a way that the total number of bad sinks decreases without significantly increasing the expected total cost.

**Lemma 11.** *Let  $\varepsilon > 0$  be arbitrary, and let  $\mathbf{d}$  be any globally  $\varepsilon$ -optimal delay vector in  $\mathcal{M}$  with  $k > 0$  bad sinks. Then there is an globally  $(2\varepsilon)$ -optimal delay vector  $\mathbf{d}'$  in  $\mathcal{M}$  with  $k - 1$  bad sinks.*

Before we prove the previous lemma, let us note that it already implies that for each  $\varepsilon > 0$  there is globally  $\varepsilon$ -optimal delay function such that all sinks are good. Indeed, a stable globally  $\varepsilon$ -optimal delay function  $\mathbf{d}^*$  can be constructed by taking any globally  $\varepsilon/2^{|S_{\text{fd}}|}$ -optimal delay function and iteratively using Lemma 11 to remove all of its at most  $|S'|$  bad sinks.

We devote the rest of this section to the proof of Lemma 11, which is perhaps the most technically involved part in the proof of Theorem 1. Fix an arbitrary globally  $\varepsilon$ -optimal delay function  $\mathbf{d}$  in  $\mathcal{M}$  and let  $K$  be any sink bad for  $\mathbf{d}$ . For any state  $s \in K$  we define the value

$$\begin{aligned} \overline{V}_{\mathcal{M}(\mathbf{d})}^s &= 1/\lambda \cdot \mathcal{R}(s) + \sum_{s' \in S'} \mathbf{P}(s, s') \cdot (E_{\mathcal{M}[s'](\mathbf{d})} + \mathcal{I}_{\mathbf{P}}(s, s')) \\ &\quad + \sum_{s'' \in S_{\text{fd}}} \sum_{s' \in S'} \mathbf{P}(s, s'') \mathbf{F}(s'', s') \cdot (E_{\mathcal{M}[s'](\mathbf{d})} + \mathcal{I}_{\mathbf{P}}(s, s'') + \mathcal{I}_{\mathbf{F}}(s'', s)). \end{aligned}$$

Intuitively the value is expected cost paid from  $s$  if exactly one exponential transition occurs in  $s$  before reaching  $S'$ . We denote the probability mass function of Poisson distribution with parameter  $\lambda > 0$  as  $f_{\text{Po}(\lambda)} : \mathbb{N}_0 \rightarrow [0, 1]$ , where  $f_{\text{Po}(\lambda)}(i) = e^{-\lambda} \cdot \frac{\lambda^i}{i!}$ . Furthermore we denote  $f_{\text{Po}(\lambda)}(\geq i) = \sum_{j=i}^{\infty} f_{\text{Po}(\lambda)}(j)$ .

**Definition 8.** *Let  $C(\mathbf{d})$  be a fdCTMC. We define a jump chain for  $C(\mathbf{d})$  as a series of random variables  $(X^{C(\mathbf{d})})_n$  where for all  $n \in \mathbb{N}_0$ ,  $X_n^{C(\mathbf{d})} : \Omega \rightarrow S$  and for a run  $\omega = (s_0, d_0)t_0 \dots$*

$$X_n^{C(\mathbf{d})}(\omega) = s_n.$$



Similarly we define a time chain for  $C(\mathbf{d})$  as a series of random variables  $(Y^{C(\mathbf{d})})_n$  where for all  $n \in \mathbb{N}_0$ ,  $Y_n^{C(\mathbf{d})} : \Omega \rightarrow \mathbb{R}_{\geq 0}$  and for a run  $\omega = (s_0, d_0)t_0 \dots$

$$Y_n^{C(\mathbf{d})}(\omega) = t_n.$$

The next technical lemma gives tight upper and lower bounds on  $E_{\mathcal{M}[s](\mathbf{d})}$  of  $s$  in bad sink using its  $\bar{V}_{\mathcal{M}(\mathbf{d})}^s$ .

**Lemma 12.** *Let  $\varepsilon > 0$  be arbitrary, and let  $\mathbf{d}$  be any globally  $\varepsilon$ -optimal delay function in  $\mathcal{M}$ . For each sink  $K$  bad for  $\mathbf{d}$ , state  $s \in K$  and for every  $d \in \{\underline{d}, \mathbf{d}(s)\}$  it holds that*

$$A(d, \mathbf{d}, s) \leq E_{\mathcal{M}[s](\mathbf{d}[s/d])} \leq A(d, \mathbf{d}, s) + B(d, \varepsilon),$$

where

$$A(d, \mathbf{d}, s) = f_{Po(\lambda d)}(0) \cdot \sum_{s' \in K} F(s, s') \cdot E_{\mathcal{M}[s'](\mathbf{d})} + f_{Po(\lambda d)}(1) \cdot \bar{V}_{\mathcal{M}(\mathbf{d})}^s$$

and

$$B(d, \varepsilon) = \lambda d \cdot f_{Po(\lambda d)}(\geq 1) \cdot \left( \text{Val}[\mathcal{M}] + \varepsilon + \max R \cdot \left( 2 + \frac{2}{\lambda} \right) \right).$$

*Proof.* We fix a bad sink  $K$ . For each state  $s \in K$  holds

$$E_{\mathcal{M}[s](\mathbf{d}[s/d])} = \sum_{s' \in S'} T(s, d)(s') \cdot E_{\mathcal{M}[s'](\mathbf{d})} + \mathbb{E}(s, d).$$

We start evaluating the right hand side. The most efficient computation of the right hand side can be done by employing uniformization approach. We build a fdCTMC  $C'[s](d)$  according to Definition 7. By computing the transient analysis of  $C'[s](d)$  at time  $d$  we can efficiently compute the  $T(s, d)(s')$  for each state  $s' \in S'$ :

$$\sum_{s' \in S'} T(s, d)(s') \cdot E_{\mathcal{M}[s'](\mathbf{d})} = \quad (16)$$

$$\begin{aligned} &= \sum_{s' \in S'} \sum_{i=0}^{\infty} \Pr_{C'[s](d)}(\text{exactly } i \text{ exponential steps with rate } \lambda \text{ happen in time } d) \\ &\quad \cdot \Pr_{C'[s](d)}(\text{reach } s' \text{ from } s \mid i \text{ exp. steps happened}) \cdot E_{\mathcal{M}[s'](\mathbf{d})} \\ &= \sum_{s' \in S'} \sum_{i=0}^{\infty} f_{Po(\lambda d)}(i) \cdot \Pr_{C'[s](d)}(\text{reach } s' \text{ from } s \mid i \text{ exp. steps happened}) \cdot E_{\mathcal{M}[s'](\mathbf{d})} \\ &= \sum_{i=0}^{\infty} f_{Po(\lambda d)}(i) \cdot \sum_{s' \in S'} \Pr_{C'[s](d)}(\text{reach } s' \text{ from } s \mid i \text{ exp. steps happened}) \cdot E_{\mathcal{M}[s'](\mathbf{d})} \\ &= \sum_{i=0}^1 f_{Po(\lambda d)}(i) \cdot \sum_{s' \in S'} \Pr_{C'[s](d)}(\text{reach } s' \text{ from } s \mid i \text{ exp. steps happened}) \cdot E_{\mathcal{M}[s'](\mathbf{d})} \\ &\quad + \sum_{i=2}^{\infty} f_{Po(\lambda d)}(i) \cdot \sum_{s' \in S'} \Pr_{C'[s](d)}(\text{reach } s' \text{ from } s \mid i \text{ exp. steps happened}) \cdot E_{\mathcal{M}[s'](\mathbf{d})} \\ &= f_{Po(\lambda d)}(0) \cdot \sum_{s' \in S} F(s, s') E_{\mathcal{M}[s'](\mathbf{d})} \quad (17) \end{aligned}$$

$$+ f_{Po(\lambda d)}(1) \cdot \left( \sum_{s' \in S'} P(s, s') \cdot E_{\mathcal{M}[s'](\mathbf{d})} + \sum_{s' \in S_{\text{fid}}} \sum_{s'' \in S'} P(s, s'') \cdot F(s'', s') \cdot E_{\mathcal{M}[s'](\mathbf{d})} \right) \quad (18)$$

$$+ \sum_{i=2}^{\infty} f_{Po(\lambda d)}(i) \cdot \sum_{s' \in S'} \Pr_{C'[s](d)}(\text{reach } s' \text{ from } s \mid i \text{ exp. steps happened}) \cdot E_{\mathcal{M}[s'](\mathbf{d})}. \quad (19)$$

We can now simply bound the  $\sum_{s' \in S'} T(s, d)(s') \cdot E_{\mathcal{M}[s'](\mathbf{d})}$  by approximating the expression (19):

$$0 \leq \sum_{i=2}^{\infty} f_{Po(\lambda d)}(i) \cdot \sum_{s' \in S'} \Pr_{C'[s](d)}(\text{reach } s' \text{ from } s \mid i \text{ exp. steps happened}) \cdot E_{\mathcal{M}[s'](\mathbf{d})} \leq \quad (20)$$

$$\begin{aligned} &\leq \sum_{i=2}^{\infty} f_{Po(\lambda d)}(i) \cdot \sum_{s' \in S'} \Pr_{C'[s](d)}(\text{reach } s' \text{ from } s \mid i \text{ exp. steps happened}) \cdot (\overline{\text{Val}}[\mathcal{M}] + \varepsilon) \\ &= \sum_{i=2}^{\infty} f_{Po(\lambda d)}(i) \cdot (\overline{\text{Val}}[\mathcal{M}] + \varepsilon) \\ &= f_{Po(\lambda d)}(\geq 2) \cdot (\overline{\text{Val}}[\mathcal{M}] + \varepsilon), \quad (21) \end{aligned}$$

where the last inequality follows from global  $\varepsilon$ -optimality of  $\mathbf{d}$ .

Let us now evaluate the  $\mathbb{E}(s, d)$ . We first define random variables  $\text{CostR}[S']$  and  $\text{CostI}[S']$  that assign to each run  $\omega = (s_0, d_0)t_0 \cdots$  the *total rate and impulse cost before*

reaching  $S'$  (in at least one transition), respectively, given by

$$CostR[S'](\omega) = \begin{cases} \sum_{i=0}^{n-1} t_i \cdot \mathcal{R}(s_i) & \text{for minimal } n > 0 \text{ such that } (s_n, d_n) \in S', \\ \infty & \text{if there is no such } n, \end{cases}$$

and

$$CostI[S'](\omega) = \begin{cases} \sum_{i=0}^{n-1} \mathcal{I}_i(\omega) & \text{for minimal } n > 0 \text{ such that } (s_n, d_n) \in S', \\ \infty & \text{if there is no such } n, \end{cases}$$

where  $\mathcal{I}_i(\omega)$  equals  $\mathcal{I}_P(s_i, s_{i+1})$  for an exp-delay transition, i.e. when  $t_i < d_i$ , and equals  $\mathcal{I}_F(s_i, s_{i+1})$  for a fixed-delay transition, i.e. when  $t_i = d_i$ . We can divide  $\mathbb{E}(s, d)$  in two parts:

$$\mathbb{E}(s, d) = E_{C[s, d]}^{Cost[S']} = E_{C'[s](d)}^{CostI[S']} + E_{C'[s](d)}^{CostR[S']} \quad (22)$$

We will similarly as above use uniformization to underestimate the expected impulse cost:

$$E_{C'[s](d)}^{CostI[S']} = \sum_{i=0}^{\infty} f_{Po(\lambda d)}(i) \cdot \sum_{s' \in S'} \Pr_{C'[s](d)}(\text{reach } s' \text{ from } s \mid i \text{ exp. steps happened}) \\ \cdot \mathbb{E}_{C'[s](d)}[CostI[S'] \mid \text{reach } s' \text{ from } s, i \text{ exp. steps happened}]$$

$$E_{C'[s](d)}^{CostI[S']} = \sum_{i=0}^1 f_{Po(\lambda d)}(i) \cdot \sum_{s' \in S'} \Pr_{C'[s](d)}(\text{reach } s' \text{ from } s \mid i \text{ exp. steps happened}) \\ \cdot \mathbb{E}_{C'[s](d)}[CostI[S'] \mid \text{reach } s' \text{ from } s, i \text{ exp. steps happened}] \quad (23)$$

$$+ \sum_{i=2}^{\infty} f_{Po(\lambda d)}(i) \cdot \sum_{s' \in S'} \Pr_{C'[s](d)}(\text{reach } s' \text{ from } s \mid i \text{ exp. steps happened}) \\ \cdot \mathbb{E}_{C'[s](d)}[CostI[S'] \mid \text{reach } s' \text{ from } s, i \text{ exp. steps happened}]. \quad (24)$$

Now we evaluate (23) and (24):

$$\sum_{i=0}^1 f_{Po(\lambda d)}(i) \cdot \sum_{s' \in S'} \Pr_{C'[s](d)}(\text{reach } s' \text{ from } s \mid i \text{ exp. steps happened}) \quad (27)$$

$$\cdot \mathbb{E}_{C'[s](d)}[CostI[S'] \mid \text{reach } s' \text{ from } s, i \text{ exp. steps happened}] = \quad (28)$$

$$= \alpha + f_{Po(\lambda d)}(0) \cdot \sum_{s' \in S'} F(s, s') \cdot \mathcal{I}_F(s, s') \\ = \alpha, \quad (29)$$

where

$$\alpha = f_{Po(\lambda d)}(1) \cdot \left( \sum_{s' \in S'} P(s, s') \cdot \mathcal{I}_P(s, s') \right. \\ \left. + \sum_{s' \in S_{\text{id}}} \sum_{s'' \in S'} P(s, s'') F(s'', s') \cdot (\mathcal{I}_P(s, s'') + \mathcal{I}_F(s'', s')) \right). \quad (30)$$

The last equation follows from the fact that all fixed-delay transitions from  $s$  go to states in  $K$  and from zero impulse cost of such transitions because  $K$  is a bad sink.

We now over and underestimate terms (25) and (26).

$$0 \leq \sum_{i=2}^{\infty} f_{Po(\lambda d)}(i) \cdot \sum_{s' \in S'} \Pr_{C'[s](d)}(\text{reach } s' \text{ from } s \mid i \text{ exp. steps happened}) \quad (31)$$

$$\cdot \mathbb{E}_{C'[s](d)}[CostI[S'] \mid \text{reach } s' \text{ from } s, i \text{ exp. steps happened}] \quad (32)$$

$$\begin{aligned} &\leq \sum_{i=2}^{\infty} f_{Po(\lambda d)}(i) \cdot \sum_{s' \in S'} \Pr_{C'[s](d)}(\text{reach } s' \text{ from } s \mid i \text{ exp. steps happened}) \cdot 2 \cdot i \cdot \max R \\ &\leq 2 \cdot \max R \cdot \sum_{i=2}^{\infty} f_{Po(\lambda d)}(i) \cdot i \\ &= 2 \cdot \max R \cdot \left( \sum_{i=0}^{\infty} f_{Po(\lambda d)}(i) \cdot i - \sum_{i=0}^1 f_{Po(\lambda d)}(i) \cdot i \right) \\ &= 2 \cdot \max R \cdot (\lambda d - f_{Po(\lambda d)}(1)). \end{aligned} \quad (33)$$

Using equations (23), (24), (25), (26), (27), (28), (29), (30), (31), (32), (33) we can under and overestimate the expected impulse cost accumulated on paths from  $s$  to  $S'$  by:

$$\alpha \leq E_{C'[s](d)}^{CostI[S']} \leq \alpha + 2 \cdot \max R \cdot (\lambda d - f_{Po(\lambda d)}(1)). \quad (34)$$

We evaluate the rate cost part of  $\mathbb{E}(s, d)$ .

$$\begin{aligned} E_{C'[s](d)}^{CostR[S']} &= \sum_{s' \in S'} (\text{expected time spent in } s' \text{ until hitting } S') \cdot \mathcal{R}(s') \\ &= (\text{expected time spent in } s \text{ until the first transition is taken}) \cdot \mathcal{R}(s) \\ &\quad + \sum_{s' \in S'} (\text{expected time spent in } s' \text{ until hitting } S' \text{ except the 1st trans.}) \cdot \mathcal{R}(s'), \end{aligned}$$

where

(expected time spent in  $s$  until the first transition is taken) =

$$\begin{aligned}
&= \int_0^d t \cdot \lambda \cdot e^{-\lambda t} dt + d \cdot e^{-\lambda d} \\
&= -d \cdot e^{-\lambda d} + \frac{1 - e^{-\lambda d}}{\lambda} + d \cdot e^{-\lambda d} \\
&= \frac{1 - e^{-\lambda d}}{\lambda} \\
&= \frac{1 - f_{Po(\lambda d)}(0)}{\lambda} \\
&= \frac{f_{Po(\lambda d)}(\geq 1)}{\lambda} \\
&= \frac{f_{Po(\lambda d)}(1)}{\lambda} + \frac{f_{Po(\lambda d)}(\geq 2)}{\lambda}
\end{aligned}$$

Thus

$$\begin{aligned}
E_{C'[s](d)}^{CostR[S']} &= (\text{expected time spent in } s \text{ until the first transition is taken}) \cdot \mathcal{R}(s) \\
&\quad + \sum_{s' \in S} (\text{expected time spent in } s' \text{ until hitting } S' \text{ except the 1st trans.}) \cdot \mathcal{R}(s'), \\
&= \left( \frac{f_{Po(\lambda d)}(1)}{\lambda} + \frac{f_{Po(\lambda d)}(\geq 2)}{\lambda} \right) \cdot \mathcal{R}(s) \\
&\quad + \sum_{s' \in S} (\text{expected time spent in } s' \text{ until hitting } S' \text{ except the 1st trans.}) \cdot \mathcal{R}(s'), \\
&\leq \left( \frac{f_{Po(\lambda d)}(1)}{\lambda} + \frac{f_{Po(\lambda d)}(\geq 2)}{\lambda} \right) \cdot \mathcal{R}(s) \\
&\quad + \sum_{s' \in S} (\text{expected time spent in } s' \text{ until hitting } S' \text{ except the 1st trans.}) \cdot \max R, \\
&\leq \left( \frac{f_{Po(\lambda d)}(1)}{\lambda} + \frac{f_{Po(\lambda d)}(\geq 2)}{\lambda} \right) \cdot \mathcal{R}(s) + \left( d - \frac{f_{Po(\lambda d)}(\geq 2)}{\lambda} \right) \cdot \max R,
\end{aligned}$$

Using above results we can easily bound from above and below the (expected rate cost accumulated on paths from  $s$  to  $S'$ ):

$$\frac{f_{Po(\lambda d)}(1)}{\lambda} \cdot \mathcal{R}(s) \leq E_{C'[s](d)}^{CostR[S']} \leq \tag{35}$$

$$\leq \frac{f_{Po(\lambda d)}(1)}{\lambda} \cdot \mathcal{R}(s) + \frac{f_{Po(\lambda d)}(\geq 2)}{\lambda} \cdot \mathcal{R}(s) + \left( d - \frac{f_{Po(\lambda d)}(\geq 2)}{\lambda} \right) \cdot \max R. \tag{36}$$

Now we have all the results needed to prove the lemma. To get the lower bound we combine and evaluate (16), (17), (18), (19), (20), (22), (30), (34), (35):

$$\begin{aligned}
& E_{\mathcal{M}[s](\mathbf{d}[s/d])} = \\
& = \sum_{s' \in S'} T(s, d)(s') \cdot E_{\mathcal{M}[s'](\mathbf{d})} + \epsilon(s, d) \\
& \geq f_{Po(\lambda d)}(0) \cdot \sum_{s' \in S} F(s, s') E_{\mathcal{M}[s'](\mathbf{d})} \\
& \quad + f_{Po(\lambda d)}(1) \cdot \left( \sum_{s' \in S'} P(s, s') \cdot E_{\mathcal{M}[s'](\mathbf{d})} + \sum_{s' \in S_{\text{fd}}} \sum_{s'' \in S'} P(s, s'') \cdot F(s'', s') \cdot E_{\mathcal{M}[s'](\mathbf{d})} \right) \\
& \quad + f_{Po(\lambda d)}(1) \\
& \quad \cdot \left( \sum_{s' \in S'} P(s, s') \cdot \mathcal{I}_P(s, s') + \sum_{s' \in S_{\text{fd}}} \sum_{s'' \in S'} P(s, s'') F(s'', s') \cdot (\mathcal{I}_P(s, s'') + \mathcal{I}_F(s'', s')) \right) \\
& \quad + \frac{f_{Po(\lambda d)}(1)}{\lambda} \cdot \mathcal{R}(s) \\
& = f_{Po(\lambda d)}(0) \cdot \sum_{s' \in S} F(s, s') E_{\mathcal{M}[s'](\mathbf{d})} + f_{Po(\lambda d)}(1) \cdot \bar{V}_{\mathcal{M}(\mathbf{d})}^s \\
& = f_{Po(\lambda d)}(0) \cdot \sum_{s' \in K} F(s, s') E_{\mathcal{M}[s'](\mathbf{d})} + f_{Po(\lambda d)}(1) \cdot \bar{V}_{\mathcal{M}(\mathbf{d})}^s = A(d, \mathbf{d}, s)
\end{aligned}$$

The second equality from bottom follows from definition of bad sink, i.e. all fixed-delay transitions from  $s$  go to states in  $K$ . For the upper bound we get the same reductions, but we need to add additional summands, see (16), (17), (18), (19), (20), (21), (22), (30), (34), (35), (36). We get the final result:

$$\begin{aligned}
E_{\mathcal{M}[s](\mathbf{d}[s/d])} & \leq f_{Po(\lambda d)}(0) \cdot \sum_{s' \in K} F(s, s') E_{\mathcal{M}[s'](\mathbf{d})} + f_{Po(\lambda d)}(1) \cdot \bar{V}_{\mathcal{M}(\mathbf{d})}^s \\
& \quad + f_{Po(\lambda d)}(\geq 2) \cdot (\overline{Val}[\mathcal{M}] + \varepsilon) + 2 \cdot \max R \cdot (\lambda d - f_{Po(\lambda d)}(1)) \\
& \quad + \frac{f_{Po(\lambda d)}(\geq 2)}{\lambda} \cdot \mathcal{R}(s) + \left( d - \frac{f_{Po(\lambda d)}(\geq 1)}{\lambda} \right) \cdot \max R \\
& \leq A(d, \mathbf{d}, s) + f_{Po(\lambda d)}(\geq 2) \cdot \left( \overline{Val}[\mathcal{M}] + \varepsilon + \frac{\max R}{\lambda} \right) \\
& \quad + (\lambda d - f_{Po(\lambda d)}(1)) \cdot \max R \cdot \left( 2 + \frac{1}{\lambda} \right) \\
& \leq A(d, \mathbf{d}, s) + \lambda d \cdot f_{Po(\lambda d)}(\geq 1) \cdot \left( \overline{Val}[\mathcal{M}] + \varepsilon + \max R \cdot \left( 2 + \frac{2}{\lambda} \right) \right) \\
& = A(d, \mathbf{d}, s) + B(d, \varepsilon),
\end{aligned}$$

where the last inequality follows from

$$\lambda d - f_{Po(\lambda d)}(1) = \lambda d(1 - e^{-\lambda d}) = \lambda d \cdot f_{Po(\lambda d)}(\geq 1) = e^{-\lambda d} \cdot \sum_{i=2}^{\infty} \frac{(\lambda d)^i}{(i-1)!} \geq f_{Po(\lambda d)}(\geq 2).$$

**Lemma 13.** Let  $\mathbf{d}$  be globally  $\varepsilon$ -optimal delay function in  $\mathcal{M}$ . Let  $K$  be a bad sink for  $\mathbf{d}$  and let  $s \in K$  be such that  $\bar{V}_{\mathcal{M}(\mathbf{d})}^s = \min_{s'' \in K} \bar{V}_{\mathcal{M}(\mathbf{d})}^{s''}$ . Then for each  $s' \in K$  it holds that

$$E_{\mathcal{M}[s'](\mathbf{d})} \geq \frac{f_{Po(\lambda \underline{d})}(1)}{(1 - f_{Po(\lambda \underline{d})}(0))} \cdot \bar{V}_{\mathcal{M}(\mathbf{d})}^s,$$

and

$$E_{\mathcal{M}[s'](\mathbf{d})} \geq V_{\mathcal{M}(\mathbf{d})}^s - \frac{f_{Po(\lambda \underline{d})}(\geq 2) \cdot (\overline{\text{Val}}[\mathcal{M}] + \varepsilon)}{f_{Po(\lambda \underline{d})}(1)}.$$

*Proof.* We fix a sink  $K$  bad for  $\mathbf{d}$  and state  $s' \in K$  with minimal value of  $E_{\mathcal{M}[s'](\mathbf{d})}$ . To obtain the lower bound on  $E_{\mathcal{M}[s'](\mathbf{d})}$  we use Lemma 12.

$$\begin{aligned} E_{\mathcal{M}[s'](\mathbf{d})} &\geq f_{Po(\lambda \mathbf{d}(s'))}(0) \cdot \sum_{s'' \in K} F(s', s'') E_{\mathcal{M}[s''](\mathbf{d})} + f_{Po(\lambda \mathbf{d}(s'))}(1) \cdot \bar{V}_{\mathcal{M}(\mathbf{d})}^{s'} \\ &\geq f_{Po(\lambda \mathbf{d}(s'))}(0) \cdot \sum_{s'' \in K} F(s', s'') E_{\mathcal{M}[s'](\mathbf{d})} + f_{Po(\lambda \mathbf{d}(s'))}(1) \cdot \bar{V}_{\mathcal{M}(\mathbf{d})}^{s'} \end{aligned} \quad (37)$$

$$\begin{aligned} &= f_{Po(\lambda \mathbf{d}(s'))}(0) \cdot E_{\mathcal{M}[s'](\mathbf{d})} + f_{Po(\lambda \mathbf{d}(s'))}(1) \cdot \bar{V}_{\mathcal{M}(\mathbf{d})}^{s'} \\ &\geq f_{Po(\lambda \mathbf{d}(s'))}(0) \cdot E_{\mathcal{M}[s'](\mathbf{d})} + f_{Po(\lambda \mathbf{d}(s'))}(1) \cdot \bar{V}_{\mathcal{M}(\mathbf{d})}^s, \end{aligned} \quad (38)$$

where the inequality (37) follows from minimality of  $E_{\mathcal{M}[s'](\mathbf{d})}$  among all states of  $s' \in K$  and the inequality (38) follows from the minimality of  $s \in K$  with respect to  $\bar{V}_{\mathcal{M}(\mathbf{d})}^s$ . We proceed with simplification of inequality (38):

$$\begin{aligned} E_{\mathcal{M}[s'](\mathbf{d})} &\geq f_{Po(\lambda \mathbf{d}(s'))}(0) \cdot E_{\mathcal{M}[s'](\mathbf{d})} + f_{Po(\lambda \mathbf{d}(s'))}(1) \cdot \bar{V}_{\mathcal{M}(\mathbf{d})}^s \\ (1 - f_{Po(\lambda \mathbf{d}(s'))}(0)) \cdot E_{\mathcal{M}[s'](\mathbf{d})} &\geq f_{Po(\lambda \mathbf{d}(s'))}(1) \cdot \bar{V}_{\mathcal{M}(\mathbf{d})}^s \\ E_{\mathcal{M}[s'](\mathbf{d})} &\geq \frac{f_{Po(\lambda \mathbf{d}(s'))}(1)}{(1 - f_{Po(\lambda \mathbf{d}(s'))}(0))} \cdot \bar{V}_{\mathcal{M}(\mathbf{d})}^s \\ &\geq \frac{f_{Po(\lambda \underline{d})}(1)}{(1 - f_{Po(\lambda \underline{d})}(0))} \cdot \bar{V}_{\mathcal{M}(\mathbf{d})}^s, \end{aligned}$$

where the last inequality can be justified as follows:  $s'$  belongs to the bad sink  $K$ , thus from definition of a bad sink it holds that  $0 < \mathbf{d}(s') \leq \underline{d}$ . Since function

$$\frac{f_{Po(\lambda \mathbf{d}(s'))}(1)}{(1 - f_{Po(\lambda \mathbf{d}(s'))}(0))} = \frac{\lambda \mathbf{d}(s') \cdot e^{-\lambda \mathbf{d}(s')}}{e^{-\lambda \mathbf{d}(s')} \cdot \sum_{i=1}^{\infty} \frac{(\lambda \mathbf{d}(s'))^i}{i!}} = \frac{1}{\sum_{i=1}^{\infty} \frac{(\lambda \mathbf{d}(s'))^{i-1}}{i!}}$$

is monotone, and from  $0 < \mathbf{d}(s') \leq \underline{d}$  we get the result. We picked the state  $s'$  in  $K$  with minimal  $E_{\mathcal{M}[s'](\mathbf{d})}$  thus for all  $s'' \in K$  it holds that

$$E_{\mathcal{M}[s''](\mathbf{d})} \geq E_{\mathcal{M}[s'](\mathbf{d})} \geq \frac{f_{Po(\lambda \underline{d})}(1)}{(1 - f_{Po(\lambda \underline{d})}(0))} \cdot \bar{V}_{\mathcal{M}(\mathbf{d})}^s.$$

Now we simplify again the inequality (38) to get the second bound:

$$\begin{aligned}
& \sum_{i=0}^{\infty} f_{Po(\lambda \mathbf{d}(s'))}(i) \cdot E_{\mathcal{M}[s'](\mathbf{d})} \geq f_{Po(\lambda \mathbf{d}(s'))}(0) \\
& \quad \cdot E_{\mathcal{M}[s'](\mathbf{d})} + f_{Po(\lambda \mathbf{d}(s'))}(1) \cdot \bar{V}_{\mathcal{M}(\mathbf{d})}^s \\
& f_{Po(\lambda \mathbf{d}(s'))}(1) \cdot E_{\mathcal{M}[s'](\mathbf{d})} + f_{Po(\lambda \mathbf{d}(s'))}(\geq 2) \cdot E_{\mathcal{M}[s'](\mathbf{d})} \geq f_{Po(\lambda \mathbf{d}(s'))}(1) \cdot \bar{V}_{\mathcal{M}(\mathbf{d})}^s \\
& E_{\mathcal{M}[s'](\mathbf{d})} + \frac{f_{Po(\lambda \mathbf{d}(s'))}(\geq 2)}{f_{Po(\lambda \mathbf{d}(s'))}(1)} \cdot E_{\mathcal{M}[s'](\mathbf{d})} \geq \bar{V}_{\mathcal{M}(\mathbf{d})}^s \\
& E_{\mathcal{M}[s'](\mathbf{d})} + \frac{f_{Po(\lambda \mathbf{d}(s'))}(\geq 2) \cdot (\bar{Val}[\mathcal{M}] + \varepsilon)}{f_{Po(\lambda \mathbf{d}(s'))}(1)} \geq \bar{V}_{\mathcal{M}(\mathbf{d})}^s \\
& E_{\mathcal{M}[s'](\mathbf{d})} \geq \bar{V}_{\mathcal{M}(\mathbf{d})}^s - \frac{f_{Po(\lambda \mathbf{d})}(\geq 2) \cdot (\bar{Val}[\mathcal{M}] + \varepsilon)}{f_{Po(\lambda \mathbf{d})}(1)},
\end{aligned}$$

where the last inequality follows again from  $0 < \mathbf{d}(s') \leq \underline{d}$  and from monotonicity of the following function

$$\frac{f_{Po(\lambda \mathbf{d}(s'))}(\geq 2)}{f_{Po(\lambda \mathbf{d}(s'))}(1)} = \frac{e^{-\lambda \mathbf{d}(s')} \cdot \sum_{i=2}^{\infty} \frac{(\lambda \mathbf{d}(s'))^i}{i!}}{\lambda \mathbf{d}(s') \cdot e^{-\lambda \mathbf{d}(s')}} = \sum_{i=2}^{\infty} \frac{(\lambda \mathbf{d}(s'))^{i-1}}{i!}.$$

Since we picked the state  $s'$  in  $K$  with minimal  $E_{\mathcal{M}[s'](\mathbf{d})}$  it holds that, for all  $s'' \in K$

$$E_{\mathcal{M}[s''](\mathbf{d})} \geq E_{\mathcal{M}[s'](\mathbf{d})} \geq \bar{V}_{\mathcal{M}(\mathbf{d})}^s - \frac{f_{Po(\lambda \mathbf{d})}(\geq 2) \cdot (\bar{Val}[\mathcal{M}] + \varepsilon)}{f_{Po(\lambda \mathbf{d})}(1)}.$$

**Lemma 14.** *Let  $\varepsilon > 0$  be arbitrary, and let  $\mathbf{d}$  be any globally  $\varepsilon$ -optimal delay function in  $\mathcal{M}$ . Next, let  $K \subseteq S'$  be sink bad for  $\mathbf{d}$ , and let  $s \in K$  be state with minimal  $\bar{V}_{\mathcal{M}(\mathbf{d})}^s$  among all states in  $K$ . Then for each state  $s' \in S'$  it holds that*

$$E_{\mathcal{M}[s'](\mathbf{d}[s/\underline{d}])} - E_{\mathcal{M}[s'](\mathbf{d})} \leq B(\underline{d}, \varepsilon),$$

where  $B(\underline{d}, \varepsilon)$  is as in Lemma 12, i.e.

$$B(\underline{d}, \varepsilon) = \lambda \underline{d} \cdot f_{Po(\lambda \underline{d})}(\geq 1) \cdot \left( \bar{Val}[\mathcal{M}] + \varepsilon + \max R \cdot \left( 2 + \frac{2}{\lambda} \right) \right).$$

*Proof.* We first prove the lemma for state  $s$ . Observe that since we start in state  $s$ , then according to delay function  $\mathbf{d}[s/\underline{d}]$  we make a change of delay only in the first visit of  $s$  and then we proceed with original delay function. Thus, using Lemma 12 we start evaluating the left hand side:



$$\begin{aligned}
& E_{\mathcal{M}[\underline{s}](\mathbf{d}[\underline{s}/\underline{d}])} - E_{\mathcal{M}[\underline{s}](\mathbf{d})} \leq \\
& \leq f_{Po(\lambda \underline{d})}(0) \cdot \sum_{s' \in K} F(s, s') E_{\mathcal{M}[\underline{s}'](\mathbf{d})} + f_{Po(\lambda \underline{d})}(1) \cdot \bar{V}_{\mathcal{M}(\mathbf{d})}^s + B(\underline{d}, \varepsilon) - E_{\mathcal{M}[\underline{s}](\mathbf{d})} \\
& \leq f_{Po(\lambda \underline{d})}(0) \cdot \sum_{s' \in K} F(s, s') E_{\mathcal{M}[\underline{s}'](\mathbf{d})} + f_{Po(\lambda \underline{d})}(1) \cdot \bar{V}_{\mathcal{M}(\mathbf{d})}^s + B(\underline{d}, \varepsilon) \\
& \quad - f_{Po(\lambda \mathbf{d}(s))}(0) \cdot \sum_{s' \in K} F(s, s') E_{\mathcal{M}[\underline{s}'](\mathbf{d})} + f_{Po(\lambda \mathbf{d}(s))}(1) \cdot \bar{V}_{\mathcal{M}(\mathbf{d})}^s \\
& = (f_{Po(\lambda \underline{d})}(0) - f_{Po(\lambda \mathbf{d}(s))}(0)) \cdot \sum_{s' \in K} F(s, s') E_{\mathcal{M}[\underline{s}'](\mathbf{d})} \\
& \quad + (f_{Po(\lambda \underline{d})}(1) - f_{Po(\lambda \mathbf{d}(s))}(1)) \cdot \bar{V}_{\mathcal{M}(\mathbf{d})}^s + B(\underline{d}, \varepsilon).
\end{aligned}$$

From the definition of a bad sink it holds that  $\underline{d} > \mathbf{d}(s)$  and thus  $f_{Po(\lambda \underline{d})}(0) - f_{Po(\lambda \mathbf{d}(s))}(0) < 0$ . From Lemma 13 it further follows that for all  $s' \in K$

$$E_{\mathcal{M}[\underline{s}'](\mathbf{d})} \geq \frac{f_{Po(\lambda \underline{d})}(1)}{(1 - f_{Po(\lambda \underline{d})}(0))} \cdot \bar{V}_{\mathcal{M}(\mathbf{d})}^s.$$

Thus

$$\begin{aligned}
& E_{\mathcal{M}[\underline{s}](\mathbf{d}[\underline{s}/\underline{d}])} - E_{\mathcal{M}[\underline{s}](\mathbf{d})} \leq \\
& \leq (f_{Po(\lambda \underline{d})}(0) - f_{Po(\lambda \mathbf{d}(s))}(0)) \cdot \sum_{s' \in K} F(s, s') E_{\mathcal{M}[\underline{s}'](\mathbf{d})} \\
& \quad + (f_{Po(\lambda \underline{d})}(1) - f_{Po(\lambda \mathbf{d}(s))}(1)) \cdot \bar{V}_{\mathcal{M}(\mathbf{d})}^s + B(\underline{d}, \varepsilon) \\
& \leq (f_{Po(\lambda \underline{d})}(0) - f_{Po(\lambda \mathbf{d}(s))}(0)) \cdot \frac{f_{Po(\lambda \underline{d})}(1)}{(1 - f_{Po(\lambda \underline{d})}(0))} \cdot \bar{V}_{\mathcal{M}(\mathbf{d})}^s \\
& \quad + (f_{Po(\lambda \underline{d})}(1) - f_{Po(\lambda \mathbf{d}(s))}(1)) \cdot \bar{V}_{\mathcal{M}(\mathbf{d})}^s + B(\underline{d}, \varepsilon) \\
& = \left( (f_{Po(\lambda \underline{d})}(0) - f_{Po(\lambda \mathbf{d}(s))}(0)) \cdot \frac{f_{Po(\lambda \underline{d})}(1)}{(1 - f_{Po(\lambda \underline{d})}(0))} + (f_{Po(\lambda \underline{d})}(1) - f_{Po(\lambda \mathbf{d}(s))}(1)) \right) \cdot \bar{V}_{\mathcal{M}(\mathbf{d})}^s \\
& \quad + B(\underline{d}, \varepsilon).
\end{aligned}$$

To get the lemma it now suffices to show that the following holds

$$\left( (f_{Po(\lambda \underline{d})}(0) - f_{Po(\lambda \mathbf{d}(s))}(0)) \cdot \frac{f_{Po(\lambda \underline{d})}(1)}{(1 - f_{Po(\lambda \underline{d})}(0))} + (f_{Po(\lambda \underline{d})}(1) - f_{Po(\lambda \mathbf{d}(s))}(1)) \right) \leq 0.$$

For better readability, we use abbreviations  $\alpha(i) = f_{Po(\lambda \underline{d})}(i)$  and  $\beta(i) = f_{Po(\lambda \mathbf{d}(s))}(i)$  for  $i \in \{0, 1\}$ . Now we are ready to show that the above inequality holds:

$$\begin{aligned}
& (\alpha(0) - \beta(0)) \cdot \frac{\alpha(1)}{1 - \alpha(0)} + \alpha(1) - \beta(1) < 0 \\
& \alpha(0)\alpha(1) - \alpha(1)\beta(0) + \alpha(1) - \beta(1) - \alpha(0)\alpha(1) + \alpha(0)\beta(1) \leq 0 \\
& -\alpha(1)\beta(0) + \alpha(1) - \beta(1) + \alpha(0)\beta(1) \leq 0 \\
& \alpha(1) - \alpha(1)\beta(0) \leq \beta(1) - \alpha(0)\beta(1) \\
& \alpha(1) \cdot (1 - \beta(0)) \leq \beta(1) \cdot (1 - \alpha(0)) \\
& \frac{\alpha(1)}{1 - \alpha(0)} \leq \frac{\beta(1)}{1 - \beta(0)} \\
& \frac{f_{Po(\lambda \underline{d})}(1)}{1 - f_{Po(\lambda \underline{d})}(0)} \leq \frac{f_{Po(\lambda \mathbf{d}(s))}(1)}{1 - f_{Po(\lambda \mathbf{d}(s))}(0)}, \\
& \frac{\lambda \underline{d} \cdot e^{-\lambda \underline{d}}}{e^{-\lambda \underline{d}} \cdot \sum_{i=1}^{\infty} \frac{(\lambda \underline{d})^i}{i!}} \leq \frac{\lambda \mathbf{d}(s) \cdot e^{-\lambda \mathbf{d}(s)}}{e^{-\lambda \mathbf{d}(s)} \cdot \sum_{i=1}^{\infty} \frac{(\lambda \mathbf{d}(s))^i}{i!}} \\
& \frac{1}{\sum_{i=1}^{\infty} \frac{(\lambda \underline{d})^{i-1}}{i!}} \leq \frac{1}{\sum_{i=1}^{\infty} \frac{(\lambda \mathbf{d}(s))^{i-1}}{i!}}
\end{aligned}$$

where the last inequality follows from definition of bad sink that  $0 < \mathbf{d}(s) < \underline{d}$ .

We finished the proof for state  $s$ . It remains to show that the increase of value for any other state  $s' \in S' \setminus \{s\}$  is smaller or equal to  $B(\underline{d}, \varepsilon)$ . For every state  $s' \in S' \setminus \{s\}$  it holds

$$\begin{aligned}
E_{\mathcal{M}[s'](\mathbf{d})} &= \Pr_{\mathcal{M}(\mathbf{d})}(X_{H_1^{[s]}}^{\mathcal{M}} = s) \cdot \mathbb{E}_{\mathcal{M}[s'](\mathbf{d})}[Cost \mid X_{H_1^{[s]}}^{\mathcal{M}} = s] \\
&\quad + (1 - \Pr_{\mathcal{M}(\mathbf{d})}(X_{H_1^{[s]}}^{\mathcal{M}} = s)) \cdot \mathbb{E}_{\mathcal{M}[s'](\mathbf{d})}[Cost \mid X_{H_1^{[s]}}^{\mathcal{M}} \neq s] \\
&= \Pr_{\mathcal{M}(\mathbf{d})}(X_{H_1^{[s]}}^{\mathcal{M}} = s) \\
&\quad \cdot ((\text{expected cost until reaching } s \text{ in } \mathcal{M}(\mathbf{d}) \mid X_{H_1^{[s]}}^{\mathcal{M}} = s) + E_{\mathcal{M}[s](\mathbf{d})}) \\
&\quad + (1 - \Pr_{\mathcal{M}(\mathbf{d})}(X_{H_1^{[s]}}^{\mathcal{M}} = s)) \cdot \mathbb{E}_{\mathcal{M}[s'](\mathbf{d})}[Cost \mid X_{H_1^{[s]}}^{\mathcal{M}} \neq s]
\end{aligned}$$

and similarly

$$\begin{aligned}
E_{\mathcal{M}[s'](\mathbf{d}[s/\underline{d}])} &= \\
&= \Pr_{\mathcal{M}(\mathbf{d})}(X_{H_1^{[s]}}^{\mathcal{M}} = s) \cdot ((\text{expected cost until reaching } s \text{ in } \mathcal{M}(\mathbf{d}) \mid X_{H_1^{[s]}}^{\mathcal{M}} = s) + E_{\mathcal{M}[s](\mathbf{d}[s/\underline{d}])}) \\
&\quad + (1 - \Pr_{\mathcal{M}(\mathbf{d})}(X_{H_1^{[s]}}^{\mathcal{M}} = s)) \cdot \mathbb{E}_{\mathcal{M}[s'](\mathbf{d}[s/\underline{d}])}[Cost \mid X_{H_1^{[s]}}^{\mathcal{M}} \neq s].
\end{aligned}$$

The last equality follows from the fact that  $\mathcal{M}(\mathbf{d}[s/\underline{d}])$  behaves exactly in the same way as  $\mathcal{M}(\mathbf{d})$  until reaching  $s$ . Thus for each state  $s' \in S' \setminus \{s\}$  we are getting

$$E_{\mathcal{M}[s'](\mathbf{d}[s/\underline{d}])} - E_{\mathcal{M}[s'](\mathbf{d})} \leq \Pr_{\mathcal{M}(\mathbf{d})}(X_{H_1^{[s]}}^{\mathcal{M}} = s) \cdot B(\underline{d}, \varepsilon) \leq B(\underline{d}, \varepsilon),$$

what completes the proof.

By Lemma 14 we have that the increase in expected cost is bounded for each state if we increase delay of state  $s$  with minimal value in first visit. We show that we have positive probability that we reach target before returning to  $s$ . First we will prove basic lemma what we will need on the way.

**Lemma 15.** *For all  $x \in \mathbb{R}_{\geq 0}$  it holds that  $e^x - 1 \leq x \cdot e^x$ .*

*Proof.* The derivation of  $e^x - 1$  is  $x \cdot e^x$ . The derivation of  $x \cdot e^x$  is  $x^2 \cdot e^x + e^x$ . Observe that for all  $x \in \mathbb{R}_{\geq 0}$  it holds that

$$\begin{aligned} x \cdot e^x &\leq x^2 \cdot e^x + e^x \\ x &\leq x^2 + 1, \end{aligned}$$

what implies that  $e^x - 1$  increases slower than  $x \cdot e^x$  on positive real numbers. Since both  $e^x - 1$  and  $x \cdot e^x$  equal to zero at  $x = 0$ , it must hold that  $e^x - 1 \leq x \cdot e^x$  for all  $x \in \mathbb{R}_{\geq 0}$ .

**Lemma 16.** *Let  $\varepsilon > 0$  be arbitrary, and let  $\mathbf{d}$  be any globally  $\varepsilon$ -optimal delay function for  $\mathcal{M}$ ,  $K \subseteq S'$  is sink bad for  $\mathbf{d}$ , and  $s \in K$  is state with minimal  $\bar{V}_{\mathcal{M}(\mathbf{d})}^s$  among states in  $K$ . It holds that*

$$p_s \geq f_{Po(\lambda \mathbf{d})}(1) \cdot \frac{\min R / (2 \cdot \lambda)}{\min R / \lambda + \bar{Val}[\mathcal{M}] + \varepsilon},$$

where  $p_s$  is the probability that goal state is reached before returning to  $s$  if run starts in  $s$ , i.e.

$$p_s = \Pr_{\mathcal{M}(\mathbf{d})}(X_n^{\mathcal{M}(\mathbf{d})} \in V', \forall l. i < l < j. X_l^{\mathcal{M}(\mathbf{d})} \neq s \mid X_i^{\mathcal{M}(\mathbf{d})} = s).$$

*Proof.* First we fix sink  $K$  bad for  $\mathbf{d}$  and  $s \in K$  that has minimal  $V_{\mathcal{M}(\mathbf{d})}$  in  $K$ . Assume we have fdCTMC  $C'(\mathbf{d})$ , where  $C'$  is same to fdCTMC from which the  $\mathcal{M}$  was created except that the initial state is  $s$ . Obviously it holds that

$$p_s = \Pr_{C'(\mathbf{d})}(X_n^{C'(\mathbf{d})} \in G, \forall l. 0 < l < j. X_l^{C'(\mathbf{d})} \neq s).$$

Thus in the remaining part of this proof we will use mostly the  $C'(\mathbf{d})$ . First we provide necessary definitions. The next property denotes the probability of reaching  $G$  before  $K$  conditioned by the fact that exactly one exponential step was done from state  $s$  in  $C'(\mathbf{d})$ :

$$p'_s = \Pr_{C'(\mathbf{d})}(X_n^{C'(\mathbf{d})} \in G, \forall l. 0 < l < j. X_l^{C'(\mathbf{d})} \notin K \mid Y_1^{C'(\mathbf{d})} < \mathbf{d}(s), X_1^{C'(\mathbf{d})} + X_2^{C'(\mathbf{d})} \geq \mathbf{d}(s)).$$

We denote by  $Cost^-$  and  $Cost^{--}$  the random variables assigning to each run  $\omega = (s_0, d_0)t_0 \dots$  the total cost after first state from  $S'$  and before reaching  $G$  and total cost after first state from  $S'$  and before reaching state from  $K$  from which an exponential transition is taken, given by

$$Cost^-(\omega) = \begin{cases} \sum_{i=j}^{n-1} (t_i \cdot \mathcal{R}(s_i) + I_i(\omega)) & \text{for minimal } n > 0 \text{ such that } s_n \in G, \\ & X_j^{C'(\mathbf{d})} \in S', \text{ and } \forall l. 0 < l < j. X_l^{C'(\mathbf{d})} \notin S', \\ \infty & \text{if there is no such } n, \end{cases}$$

and

$$Cost^{--}(\omega) = \begin{cases} \sum_{i=j}^{n-1} (t_i \cdot \mathcal{R}(s_i) + \mathcal{I}_i(\omega)) & \text{for minimal } n > 0 \text{ such that } s_n \in K, Y_n^{C'(\mathbf{d})} < \mathbf{d}(s_n), \\ & X_j^{C'(\mathbf{d})} \in S', \text{ and } \forall l. 0 < l < j. X_l^{C'(\mathbf{d})} \notin S', \\ \infty & \text{if there is no such } n, \end{cases}$$

respectively, where  $\mathcal{I}_i(\omega)$  equals  $\mathcal{I}_P(s_i, s_{i+1})$  for an exp-delay transition, i.e. when  $t_i < d_i$ , and equals  $\mathcal{I}_F(s_i, s_{i+1})$  for a fixed-delay transition, i.e. when  $t_i = d_i$ .

Now we evaluate the following basic property:

$$\begin{aligned} V_{\mathcal{M}(\mathbf{d})}^s &\geq \\ &\geq p'_s \cdot \mathbb{E}_{C'(\mathbf{d})}[Cost^- \mid \text{reach } G \text{ before } K \text{ and } Y_1^{C'(\mathbf{d})} < \mathbf{d}(s), X_1^{C'(\mathbf{d})} + X_2^{C'(\mathbf{d})} \geq \mathbf{d}(s)] \\ &\quad + (1 - p'_s) \cdot \mathbb{E}_{C'(\mathbf{d})}[Cost^- \mid \text{reach } K \text{ before } G \text{ and } Y_1^{C'(\mathbf{d})} < \mathbf{d}(s), X_1^{C'(\mathbf{d})} + X_2^{C'(\mathbf{d})} \geq \mathbf{d}(s)] \\ &\geq (1 - p'_s) \cdot \mathbb{E}_{C'(\mathbf{d})}[Cost^- \mid \text{reach } K \text{ before } G \text{ and } Y_1^{C'(\mathbf{d})} < \mathbf{d}(s), X_1^{C'(\mathbf{d})} + X_2^{C'(\mathbf{d})} \geq \mathbf{d}(s)] \\ &= (1 - p'_s) \cdot \left( c + \sum_{s' \in K} \pi(s') \cdot E_{\mathcal{M}[s'](\mathbf{d})} \right) \\ &= c + \sum_{s' \in K} \pi(s') \cdot E_{\mathcal{M}[s'](\mathbf{d})} - p'_s \cdot \left( c + \sum_{s' \in K} \pi(s') \cdot E_{\mathcal{M}[s'](\mathbf{d})} \right) \\ &\geq c + \sum_{s' \in K} \pi(s') \cdot \left( V_{\mathcal{M}(\mathbf{d})}^s - \frac{f_{Po(\lambda \underline{d})}(\geq 2) \cdot (\overline{Val}[\mathcal{M}] + \varepsilon)}{f_{Po(\lambda \underline{d})}(1)} \right) \\ &\quad - p'_s \cdot \left( c + \sum_{s' \in K} \pi(s') \cdot (\overline{Val}[\mathcal{M}] + \varepsilon) \right) \\ &= c + V_{\mathcal{M}(\mathbf{d})}^s - \frac{f_{Po(\lambda \underline{d})}(\geq 2) \cdot (\overline{Val}[\mathcal{M}] + \varepsilon)}{f_{Po(\lambda \underline{d})}(1)} - p'_s \cdot (c + \overline{Val}[\mathcal{M}] + \varepsilon) \end{aligned}$$

where

$$c = \mathbb{E}_{C'(\mathbf{d})}[Cost^{--} \mid \text{reach } K \text{ before } G \text{ and } Y_1^{C'(\mathbf{d})} < \mathbf{d}(s), X_1^{C'(\mathbf{d})} + X_2^{C'(\mathbf{d})} \geq \mathbf{d}(s)],$$

and for  $s' \in K$

$$\begin{aligned} \pi(s') &= \Pr_{C'(\mathbf{d})}(\text{leaving from state } s' \text{ after returning to } K \\ &\quad \mid \text{reach } K \text{ before } G \text{ and } Y_1^{C'(\mathbf{d})} < \mathbf{d}(s), X_1^{C'(\mathbf{d})} + X_2^{C'(\mathbf{d})} \geq \mathbf{d}(s)). \end{aligned}$$

The last inequality follows from Lemma 13. By further reduction we are getting:

$$\begin{aligned}
V_{\mathcal{M}(\mathbf{d})}^s &\geq c + V_{\mathcal{M}(\mathbf{d})}^s - \frac{f_{Po(\lambda \underline{d})}(\geq 2) \cdot (\overline{Val}[\mathcal{M}] + \varepsilon)}{f_{Po(\lambda \underline{d})}(1)} - p'_s \cdot (c + \overline{Val}[\mathcal{M}] + \varepsilon) \\
p'_s \cdot (c + \overline{Val}[\mathcal{M}] + \varepsilon) &\geq c - \frac{f_{Po(\lambda \underline{d})}(\geq 2) \cdot (\overline{Val}[\mathcal{M}] + \varepsilon)}{f_{Po(\lambda \underline{d})}(1)} \\
p'_s &\geq \frac{c - \frac{f_{Po(\lambda \underline{d})}(\geq 2) \cdot (\overline{Val}[\mathcal{M}] + \varepsilon)}{f_{Po(\lambda \underline{d})}(1)}}{c + \overline{Val}[\mathcal{M}] + \varepsilon}.
\end{aligned}$$

Observe that  $c$  is actually a cost accumulated until getting to state in  $K$  from which an exponential transition is taken. We know that exponential transition is memoryless, thus the expected time until exponential transition occurs is  $1/\lambda$ . Thus it holds that  $c \geq \min R/\lambda - \max R \cdot \underline{d}$ . From conditional probability and plugging in the bound for  $c$  we get:

$$\begin{aligned}
p_s &\geq \Pr_{C'(\mathbf{d})}(Y_1^{C'(\mathbf{d})} < \mathbf{d}(s), X_1^{C'(\mathbf{d})} + X_2^{C'(\mathbf{d})} \geq \mathbf{d}(s)) \cdot p'_s \\
&\geq f_{Po(\lambda \underline{d})}(1) \cdot \frac{\min R/\lambda - \max R \cdot \underline{d} - \frac{f_{Po(\lambda \underline{d})}(\geq 2) \cdot (\overline{Val}[\mathcal{M}] + \varepsilon)}{f_{Po(\lambda \underline{d})}(1)}}{\min R/\lambda - \max R \cdot \underline{d} + \overline{Val}[\mathcal{M}] + \varepsilon} \quad (39)
\end{aligned}$$

We now simplify the equation (39) by inserting the value for  $\underline{d}$  in it. We first evaluate following:

$$\begin{aligned}
\frac{\min R}{4 \cdot \lambda} - \max R \cdot \underline{d} &= \frac{\min R}{4 \cdot \lambda} - \max R \cdot \frac{\varepsilon \cdot \min R}{4 \cdot \lambda^2 \cdot (\overline{Val}[\mathcal{M}] + \varepsilon + \max R \cdot (2 + 2/\lambda))^2} \\
&\geq \frac{\min R}{4 \cdot \lambda} - \max R \cdot \frac{\varepsilon \cdot \min R}{4 \cdot \lambda^2 \cdot (\varepsilon + \max R/\lambda)^2} \\
&\geq \frac{\min R}{4 \cdot \lambda} - \max R \cdot \frac{\varepsilon \cdot \min R}{4 \cdot \lambda^2 \cdot (2 \cdot \varepsilon \cdot \max R/\lambda)} \\
&\geq 0, \quad (40)
\end{aligned}$$

Now we bound the term  $\frac{f_{Po(\lambda \underline{d})}(\geq 2) \cdot (\overline{Val}[\mathcal{M}] + \varepsilon)}{f_{Po(\lambda \underline{d})}(1)}$  from equation (39):

$$\begin{aligned}
\frac{f_{Po(\lambda \underline{d})}(\geq 2) \cdot (\overline{Val}[\mathcal{M}] + \varepsilon)}{f_{Po(\lambda \underline{d})}(1)} &= \frac{(1 - e^{-\lambda \underline{d}} - \lambda \underline{d} \cdot e^{-\lambda \underline{d}}) \cdot (\overline{Val}[\mathcal{M}] + \varepsilon)}{\lambda \underline{d} \cdot e^{-\lambda \underline{d}}} \\
&= \frac{(e^{\lambda \underline{d}} - 1 - \lambda \underline{d}) \cdot (\overline{Val}[\mathcal{M}] + \varepsilon)}{\lambda \underline{d}} \quad (41)
\end{aligned}$$

$$\begin{aligned}
&\leq \frac{(\lambda \underline{d} \cdot e^{\lambda \underline{d}} - \lambda \underline{d}) \cdot (\overline{Val}[\mathcal{M}] + \varepsilon)}{\lambda \underline{d}} \\
&= (e^{\lambda \underline{d}} - 1) \cdot (\overline{Val}[\mathcal{M}] + \varepsilon) \quad (42)
\end{aligned}$$

We use Lemma 15 in the inequality (41). In inequality (42) we plug in the definition of  $\underline{d}$ :

$$\begin{aligned}
\frac{f_{Po(\lambda \underline{d})}(\geq 2) \cdot (\overline{Val}[\mathcal{M}] + \varepsilon)}{f_{Po(\lambda \underline{d})}(1)} &\leq (e^{\lambda \underline{d}} - 1) \cdot (\overline{Val}[\mathcal{M}] + \varepsilon) \\
&= \frac{\varepsilon \cdot \min R}{4 \cdot \lambda \cdot (\overline{Val}[\mathcal{M}] + \varepsilon + \max R \cdot (2 + 2/\lambda))^2} \cdot (\overline{Val}[\mathcal{M}] + \varepsilon) \\
&\leq \frac{\varepsilon \cdot \min R}{4 \cdot \lambda \cdot (\overline{Val}[\mathcal{M}] + \varepsilon)^2} \cdot (\overline{Val}[\mathcal{M}] + \varepsilon) \\
&= \frac{\varepsilon \cdot \min R}{4 \cdot \lambda \cdot (\overline{Val}[\mathcal{M}] + \varepsilon)} \\
&\leq \frac{\min R}{4 \cdot \lambda}, \tag{43}
\end{aligned}$$

Finally we plug inequalities (40) and (43) into inequality (39) to finish the proof

$$\begin{aligned}
p_s &\geq f_{Po(\lambda \underline{d})}(1) \cdot \frac{\min R/\lambda - \max R \cdot \underline{d} - \frac{f_{Po(\lambda \underline{d})}(\geq 2) \cdot (\overline{Val}[\mathcal{M}] + \varepsilon)}{f_{Po(\lambda \underline{d})}(1)}}{\min R/\lambda - \max R \cdot \underline{d} + \overline{Val}[\mathcal{M}] + \varepsilon} \\
&\geq f_{Po(\lambda \underline{d})}(1) \cdot \frac{\min R/(2 \cdot \lambda)}{\min R/\lambda - \max R \cdot \underline{d} + \overline{Val}[\mathcal{M}] + \varepsilon} \\
&\geq f_{Po(\lambda \underline{d})}(1) \cdot \frac{\min R/(2 \cdot \lambda)}{\min R/\lambda + \overline{Val}[\mathcal{M}] + \varepsilon}.
\end{aligned}$$

Now we have all necessary to prove Lemma 11.

**Lemma 11.** Let  $\varepsilon > 0$  be arbitrary, and let  $\mathbf{d}$  be any globally  $\varepsilon$ -optimal delay function in  $\mathcal{M}$  with  $k > 0$  bad sinks. Then there is an globally  $2\varepsilon$ -optimal delay function  $\mathbf{d}'$  in  $\mathcal{M}$  with  $k - 1$  bad sinks.

*Proof.* First we fix sink  $K$  bad for  $\mathbf{d}$  and  $s \in K$  that has minimal  $V_{\mathcal{M}(\mathbf{d})}$  in  $K$ . We will get rid of  $K$  by inflating delay of  $s$  to  $\underline{d}$ . But we will not do it in one turn but by gradually changing  $\mathbf{d}$  to  $\mathbf{d}[s/\underline{d}/1]$ ,  $\mathbf{d}[s/\underline{d}/2]$ ,  $\dots$  and showing that we can bound the change in value of any state. First we will provide the necessary random variable: For each  $n \in \mathbb{N}$  we denote by  $\overline{Cost}_n^s$  the random variable assigning to each run  $s_0, a_1 s_1 \dots$  the *total cost before  $n$ th reach of  $s$* , given by

$$\overline{Cost}_n^s(v_0 a_1 v_1 a_2 v_2 \dots) = \sum_{i=0}^{H_n^{[s]}} \epsilon(v_i, a_{i+1}).$$

Now for each  $s' \in S'$  and  $i \in \mathbb{N}$  holds

$$\begin{aligned}
E_{\mathcal{M}[s'](\mathbf{d}[s/\underline{d}/i])}(X_{H_i^{[s]}}^{\mathcal{M}} = s) &\cdot E_{\mathcal{M}[s'](\mathbf{d}[s/\underline{d}/i])}[Cost \mid X_{H_i^{[s]}}^{\mathcal{M}} = s] \\
&+ \Pr_{\mathcal{M}[s'](\mathbf{d}[s/\underline{d}/i])}(X_{H_i^{[s]}}^{\mathcal{M}} \neq s) \cdot E_{\mathcal{M}[s'](\mathbf{d}[s/\underline{d}/i])}[Cost \mid X_{H_i^{[s]}}^{\mathcal{M}} \neq s] \\
&= \Pr_{\mathcal{M}[s'](\mathbf{d}[s/\underline{d}/i])}(X_{H_i^{[s]}}^{\mathcal{M}} = s) \cdot (E_{\mathcal{M}[s'](\mathbf{d}[s/\underline{d}/i])}[\overline{Cost}_i \mid X_{H_i^{[s]}}^{\mathcal{M}} \neq s] + E_{\mathcal{M}[s](\mathbf{d})}) \\
&+ \Pr_{\mathcal{M}[s'](\mathbf{d}[s/\underline{d}/i])}(X_{H_i^{[s]}}^{\mathcal{M}} \neq s) \cdot E_{\mathcal{M}[s'](\mathbf{d}[s/\underline{d}/i])}[Cost \mid X_{H_i^{[s]}}^{\mathcal{M}} \neq s].
\end{aligned}$$

Similarly we get

$$\begin{aligned}
E_{\mathcal{M}[\mathbf{d}[s/\underline{d}/i+1])} &= \Pr_{\mathcal{M}[\mathbf{d}[s/\underline{d}/i])}(X_{H_i^{(s)}}^{\mathcal{M}} = s) \cdot E_{\mathcal{M}[\mathbf{d}[s/\underline{d}/i+1])}[Cost \mid X_{H_i^{(s)}}^{\mathcal{M}} = s] \\
&\quad + \Pr_{\mathcal{M}[\mathbf{d}[s/\underline{d}/i])}(X_{H_i^{(s)}}^{\mathcal{M}} \neq s) \cdot E_{\mathcal{M}[\mathbf{d}[s/\underline{d}/i+1])}[Cost \mid X_{H_i^{(s)}}^{\mathcal{M}} \neq s] \\
&= \Pr_{\mathcal{M}[\mathbf{d}[s/\underline{d}/i])}(X_{H_i^{(s)}}^{\mathcal{M}} = s) \cdot E_{\mathcal{M}[\mathbf{d}[s/\underline{d}/i+1])}[Cost \mid X_{H_i^{(s)}}^{\mathcal{M}} = s] \\
&\quad + \Pr_{\mathcal{M}[\mathbf{d}[s/\underline{d}/i])}(X_{H_i^{(s)}}^{\mathcal{M}} \neq s) \cdot E_{\mathcal{M}[\mathbf{d}[s/\underline{d}/i])}[Cost \mid X_{H_i^{(s)}}^{\mathcal{M}} \neq s] \\
&= \Pr_{\mathcal{M}[\mathbf{d}[s/\underline{d}/i])}(X_{H_i^{(s)}}^{\mathcal{M}} = s) \\
&\quad \cdot (E_{\mathcal{M}[\mathbf{d}[s/\underline{d}/i+1])}[\overline{Cost}_i \mid X_{H_i^{(s)}}^{\mathcal{M}} \neq s] + E_{\mathcal{M}[\mathbf{d}[s/\underline{d}/1])}) \\
&\quad + \Pr_{\mathcal{M}[\mathbf{d}[s/\underline{d}/i])}(X_{H_i^{(s)}}^{\mathcal{M}} \neq s) \cdot E_{\mathcal{M}[\mathbf{d}[s/\underline{d}/i])}[Cost \mid X_{H_i^{(s)}}^{\mathcal{M}} \neq s] \\
&= \Pr_{\mathcal{M}[\mathbf{d}[s/\underline{d}/i])}(X_{H_i^{(s)}}^{\mathcal{M}} = s) \\
&\quad \cdot (E_{\mathcal{M}[\mathbf{d}[s/\underline{d}/i])}[\overline{Cost}_i \mid X_{H_i^{(s)}}^{\mathcal{M}} \neq s] + E_{\mathcal{M}[\mathbf{d}[s/\underline{d}/1])}) \\
&\quad + \Pr_{\mathcal{M}[\mathbf{d}[s/\underline{d}/i])}(X_{H_i^{(s)}}^{\mathcal{M}} \neq s) \cdot E_{\mathcal{M}[\mathbf{d}[s/\underline{d}/i])}[Cost \mid X_{H_i^{(s)}}^{\mathcal{M}} \neq s].
\end{aligned}$$

Now subtracting to find the increase of expected cost we get

$$\begin{aligned}
E_{\mathcal{M}[\mathbf{d}[s/\underline{d}/i+1])}^{s'} - E_{\mathcal{M}[\mathbf{d}[s/\underline{d}/i])}^{s'} &= \Pr_{\mathcal{M}[\mathbf{d}[s/\underline{d}/i])}(X_{H_i^{(s)}}^{\mathcal{M}} = s) \cdot (E_{\mathcal{M}[\mathbf{d}]} - E_{\mathcal{M}[\mathbf{d}[s/\underline{d}/1])}) \\
&= (1 - p_s)^i \cdot (E_{\mathcal{M}[\mathbf{d}]} - E_{\mathcal{M}[\mathbf{d}[s/\underline{d}/1])}),
\end{aligned} \tag{44}$$

$$= (1 - p_s)^i \cdot (E_{\mathcal{M}[\mathbf{d}]} - E_{\mathcal{M}[\mathbf{d}[s/\underline{d}/1])}), \tag{45}$$

where in last inequality we used Lemma 16. Obviously for  $i = 0$  we get

$$\begin{aligned}
E_{\mathcal{M}[\mathbf{d}[s/\underline{d}/i+1])} - E_{\mathcal{M}[\mathbf{d}[s/\underline{d}/i])} &= E_{\mathcal{M}[\mathbf{d}[s/\underline{d}/1])} - E_{\mathcal{M}[\mathbf{d}]} \\
&= (1 - p_s)^0 \cdot (E_{\mathcal{M}[\mathbf{d}]} - E_{\mathcal{M}[\mathbf{d}[s/\underline{d}/1])}).
\end{aligned} \tag{46}$$

Finally putting together (44), (45), (46), Lemma 16, Lemma 14 and summing up the geometric series, we get the final result

$$\begin{aligned}
E_{\mathcal{M}[s'](\mathbf{d}[s/\underline{d}/\infty])} - E_{\mathcal{M}[s'](\mathbf{d})} &= \sum_{i=0}^{\infty} (1 - p_s)^i \cdot (E_{\mathcal{M}[s](\mathbf{d})} - E_{\mathcal{M}[s](\mathbf{d}[s/\underline{d}/1])}) \\
&= 1/p_s \cdot (E_{\mathcal{M}[s](\mathbf{d})} - E_{\mathcal{M}[s](\mathbf{d}[s/\underline{d}/1])}) \\
&\leq \frac{(\min R/\lambda + \overline{Val}[\mathcal{M}] + \varepsilon) \cdot \lambda}{2 \cdot \min R \cdot f_{Po(\lambda \underline{d})}(1)} \\
&\quad \cdot \lambda \underline{d} \cdot f_{Po(\lambda \underline{d})}(\geq 1) \cdot (\overline{Val}[\mathcal{M}] + \varepsilon + \max R \cdot (2 + \frac{2}{\lambda})) \\
&\leq \frac{(\max R/\lambda + \overline{Val}[\mathcal{M}] + \varepsilon)}{2 \cdot \min R \cdot \lambda \underline{d} \cdot e^{-\lambda \underline{d}}} \\
&\quad \cdot \lambda^2 \underline{d} \cdot (1 - e^{-\lambda \underline{d}}) \cdot (\overline{Val}[\mathcal{M}] + \varepsilon + \max R \cdot (2 + \frac{2}{\lambda})) \\
&\leq \frac{(e^{\lambda \underline{d}} - 1) \cdot \lambda \cdot (\overline{Val}[\mathcal{M}] + \varepsilon + \max R \cdot (2 + \frac{2}{\lambda}))^2}{2 \cdot \min R}
\end{aligned}$$

Finally we plug in the definition of  $\underline{d}$

$$e^{\lambda \underline{d}} = 1 + \frac{\varepsilon \cdot \min R}{4 \cdot \lambda \cdot (\overline{Val}[\mathcal{M}] + \varepsilon + \max R \cdot (2 + 2/\lambda))^2}$$

to the last inequality and we get:

$$\begin{aligned}
E_{\mathcal{M}(\mathbf{d}[s/\underline{d}/\infty])} - E_{\mathcal{M}[s'](\mathbf{d})} &\leq \frac{(e^{\lambda \underline{d}} - 1) \cdot \lambda \cdot (\overline{Val}[\mathcal{M}] + \varepsilon + \max R \cdot (2 + \frac{2}{\lambda}))^2}{2 \cdot \min R} \\
&\leq \frac{\varepsilon}{8}.
\end{aligned}$$

### C.7 Proof of Proposition 3

**Proposition 3.** For  $N$  from Lemma 4,  $D_1$  and  $D_2$  from Lemma 3, it holds that

$$\left| Val[\mathcal{M}] - Val[\mathcal{M}, D(\delta, \bar{d})] \right| \leq \frac{\varepsilon}{2}$$

$$\text{where } \delta := \frac{\alpha}{D_1}, \bar{d} := |\log(\alpha)| \cdot D_2 \cdot (\overline{Val}[\mathcal{M}] + \varepsilon), \alpha := \frac{\varepsilon^2}{64N \cdot |S'| \cdot (1 + \overline{Val}[\mathcal{M}])^2}.$$

*Proof.* Let  $\mathbf{d}'$  be the globally  $\varepsilon/4$ -optimal delay function from Lemma 4 (obtained by putting  $\varepsilon' = \varepsilon/2$ ). From the choice of parameters  $\delta$  and  $\bar{d}$  and from Lemma 3 it follows that



$D(\delta, \bar{d})$  contains a function  $\mathbf{d}$  that is  $\alpha$ -bounded by  $\mathbf{d}'$ . From Lemma 2 it follows that the cost incurred by  $\mathbf{d}$  (from any initial state) differs from the cost incurred by  $\mathbf{d}'$  by at most

$$\begin{aligned} & 2\alpha \frac{N \cdot \text{Bound}[\mathbb{E}, \mathbf{d}']}{\varepsilon} (1 + \text{Bound}[\mathbb{E}, \mathbf{d}'] |S'|) \\ & \leq \frac{8\varepsilon \cdot (1 + \varepsilon + \overline{\text{Val}}[\mathcal{M}])}{64(1 + \overline{\text{Val}}[\mathcal{M}])} \leq \frac{\varepsilon}{8} (1 + \varepsilon) \leq \varepsilon/4. \end{aligned}$$

Hence,  $\mathbf{d} \in D(\delta, \bar{d})$  is  $\varepsilon/2$ -optimal in every state  $\mathcal{M}$ , from which the Proposition follows. (Note that  $\text{Val}[\mathcal{M}, D(\delta, \bar{d})] \geq \text{Val}[\mathcal{M}]$ .)

### C.8 Proof of Lemma 5

**Lemma 5.** There is a number  $M \in \exp(\|C\|^{O(1)})$  computable in time polynomial in  $\|C\|$  such that  $\overline{\text{Val}}[\mathcal{M}] \leq M$ .

*Proof.* We simply choose a appropriate delay function and compute the overestimation of expected cost until reaching goal states from any state  $s \in S'$ . We pick a delay function  $\mathbf{d}$  such that for each  $s \in S'$  it holds that  $\mathbf{d}(s) = \frac{|S_{\text{fd}}|}{\lambda}$ .

We first bound the  $\mathbb{E}(s, \mathbf{d}(s))$  for each state  $s \in S'$ . Obviously for  $s \in S' \setminus S_{\text{fd}}$  the expected impulse cost paid in one step is  $\max R$  and expected rate cost paid in one step is  $\max R/\lambda$ , thus together  $\mathbb{E}(s, \mathbf{d}(s)) \leq (1 + 1/\lambda) \cdot \max R$ . For state  $s \in S' \cap S_{\text{fd}}$  the expected number of exponential steps until reaching  $S'$  in time  $\mathbf{d}(s) = \frac{|S_{\text{fd}}|}{\lambda}$  is  $\mathbf{d}(s) \cdot \lambda = |S_{\text{fd}}|$  (employing mean of Poisson distribution). It is possible that one extra step is taken due to fixed-delay transition, thus the expected impulse cost is bounded by  $(|S_{\text{fd}}| + 1) \cdot \max R$ . Finally the expected rate cost paid until state from  $S'$  is reached from any  $S' \cap S_{\text{fd}}$  can be bounded from above by  $\mathbf{d}(s) \cdot \max R = \frac{|S_{\text{fd}}|}{\lambda} \cdot \max R$ , what is maximum time until  $S'$  are reached  $\mathbf{d}(s) = \frac{|S_{\text{fd}}|}{\lambda}$  (from definition of  $\mathbb{E}$  after firing of fixed-delay transition a state from  $S'$  is reached) times maximum rate cost bounded by  $\max R$ . Altogether we have:

$$\begin{aligned} \mathbb{E}(s, \mathbf{d}(s)) & \leq \max\{(1 + 1/\lambda) \cdot \max R, |S_{\text{fd}}|/\lambda \cdot \max R + (|S_{\text{fd}}| + 1) \cdot \max R\} \\ & \leq (|S_{\text{fd}}|/\lambda + |S_{\text{fd}}| + 1) \cdot \max R. \end{aligned}$$

Now we bound the minimal branching probability of  $T(s, |S_{\text{fd}}|/\lambda)$ . We use again uniformization method to compute the transient analysis of  $C'[s](\mathbf{d}(s))$  to evaluate  $T(s, |S_{\text{fd}}|/\lambda)$ . For all  $s \in S' \cap S_{\text{fd}}$  and  $s' \in S'$  that is reachable from  $\bar{s}$  in  $C'[s](\mathbf{d}(s))$  ( $\bar{s}$  is artificial initial state in  $C'[s](\mathbf{d}(s))$  corresponding to state  $s$ ) it holds that

$$\begin{aligned} T(s, |S_{\text{fd}}|/\lambda)(s') & = e^{-|S_{\text{fd}}|} \sum_{i=0}^{\infty} P^i(\bar{s}, s') \cdot \frac{|S_{\text{fd}}|^i}{i!} \\ & \geq e^{-|S_{\text{fd}}|} \sum_{i=0}^{|S_{\text{fd}}|} P^i(\bar{s}, s') \cdot \frac{|S_{\text{fd}}|^i}{i!} \\ & \geq e^{-|S_{\text{fd}}|} \min_{i \in \{0, \dots, |S_{\text{fd}}|\}} \min P^{|S_{\text{fd}}|} \cdot \frac{|S_{\text{fd}}|^i}{i!} \\ & \geq e^{-|S_{\text{fd}}|} \min P^{|S_{\text{fd}}|} \\ & = (\min P/e)^{|S_{\text{fd}}|}, \end{aligned} \tag{47}$$

where (47) holds because the minimal positive branching probability in  $C'[s](\mathbf{d}(s))$  is  $\min P$  and each state, if it is reachable in  $C'[s](\mathbf{d}(s))$  from  $\bar{s}$  it has to be reached by path of length at most  $|S_{\text{fd}}|$ . Thus the minimum discrete probability to reach any reachable state is at least  $\min P^{|S_{\text{fd}}|}$ . This probability is at least once multiplied by a factor  $\frac{|S_{\text{fd}}|^i}{i^i}$  for appropriate  $i$  thus we pick the smallest factor to get the correct lower bound on probability of moving from  $S' \cap S_{\text{fd}}$ . The probability of moving from  $S' \setminus S_{\text{fd}}$  is simply  $\min P$  from definition of  $T(s, |S_{\text{fd}}|/\lambda)$ . Since  $(\min P/e)^{|S_{\text{fd}}|} \leq \min P$  it holds that for all  $s, s' \in S'$  such that  $T(s, |S_{\text{fd}}|/\lambda)(s') > 0$  implies

$$T(s, |S_{\text{fd}}|/\lambda)(s') \geq (\min P/e)^{|S_{\text{fd}}|}.$$

Now we provide a lower bound on probability of reaching  $G$  from any  $s \in S'$ :

$$\begin{aligned} \Pr_{\mathcal{M}(\mathbf{d})}(\text{reach } G \text{ from } s \text{ in } |S'| \text{ steps}) &\geq (\min \{T(s', |S_{\text{fd}}|/\lambda)(s'') > 0 \mid s', s'' \in S'\})^{|S'|} \\ &\geq ((\min P/e)^{|S_{\text{fd}}|})^{|S'|} \\ &\geq (\min P/e)^{|S|^2}. \end{aligned}$$

Finally we bound the expected cost using Bernoulli trials, i.e. for all  $s \in S'$  it holds that

$$\begin{aligned} E_{\mathcal{M}[s](\mathbf{d})}^{\infty} &\leq \sum_{i=1}^{\infty} i \cdot |S'| \cdot \max_{s'' \in S'} \in(s'', \mathbf{d}(s'')) \cdot (1 - pr)^{i-1} \cdot pr \\ &= \frac{|S'| \cdot \max_{s'' \in S'} \in(s'', \mathbf{d}(s''))}{pr} \\ &\leq \frac{|S'| \cdot (|S_{\text{fd}}| + 1 + |S_{\text{fd}}|/\lambda) \cdot \max R}{(\min P/e)^{|S|^2}}, \end{aligned} \tag{48}$$

where  $pr = \min_{s'' \in S'} \Pr_{\mathcal{M}(\mathbf{d})}(\text{reach } G \text{ from } s \text{ in } |S \setminus S_{\text{fd}}| \text{ steps})$ . The number  $M$  can be taken to be the right hand side of (48).

### C.9 Proof of Proposition 4

**Proposition 4.** Let  $\varepsilon > 0$  and fix  $\kappa = (\varepsilon \cdot \delta \cdot \min R)/(2 \cdot |S'| \cdot (1 + \overline{\text{Val}}[\mathcal{M}])^2)$ , where  $\min R$  is a minimal cost rate in  $C$ . Then it holds

$$\left| \text{Val}[\mathcal{M}, D(\delta, \bar{d})] - \text{Val}[\mathcal{M}^*] \right| \leq \frac{\varepsilon}{2}.$$

Let  $\mathbf{d}$  be any delay function in  $D(\delta, \bar{d})$ . Note that  $\mathbf{d}$  can be taken as a delay function both in  $\mathcal{M}$  and  $\mathcal{M}^*$ . To distinguish these two contexts we denote this function by  $\mathbf{d}_\kappa$  when viewing it as a strategy in  $\mathcal{M}_\kappa$ . These “two” functions  $\kappa$ -bound one another, i.e. for all states  $s$  and  $t$  it holds

1.  $|T_\kappa(s, \mathbf{d}_\kappa(s))(t) - T(s, \mathbf{d}(s))(t)| \leq \kappa$  and
2.  $|\in_\kappa(s, \mathbf{d}_\kappa(s)) - \in(s, \mathbf{d}(s))| \leq \kappa$ ,

with the qualitative properties of transitions being preserved. Hence, we can use exactly the same arguments as in the proof of Lemma 2 to show that the difference between costs incurred by these strategies is bounded by these two expressions:

$$E_1 = 2\kappa \cdot \text{Bound}[\#, \mathbf{d}] (1 + \text{Bound}[\epsilon, \mathbf{d}] \cdot |S'|) \quad (49)$$

$$E_2 = 2\kappa \cdot \text{Bound}[\#, \mathbf{d}_\kappa] (1 + \text{Bound}[\epsilon_\kappa, \mathbf{d}_\kappa] \cdot |S'|). \quad (50)$$

Now let  $\mathbf{d} \in D(\delta, \bar{d})$  be any  $(\varepsilon/8)$ -optimal strategy in  $\mathcal{M}$ . Then  $\text{Bound}[\epsilon, \mathbf{d}] \leq \overline{\text{Val}}[\mathcal{M}] + \varepsilon/8$ . Moreover, since minimal delay used in  $\mathbf{d}$  is equal to  $\delta$ , the minimal one-step cost incurred by  $\mathbf{d}$  is equal to  $((1 - e^{-\lambda\delta})/\lambda) \cdot \min R$  (this is because the average time elapsed before either firing an exponential transition or the timeout  $\delta$  running out is  $(1 - e^{-\lambda\delta})/\lambda$ ). Hence, on average,  $\mathbf{d}$  performs at most  $\frac{(\overline{\text{Val}}[\mathcal{M}] + \varepsilon) \cdot \lambda}{(1 - e^{-\lambda\delta}) \min R} \leq \frac{2(\overline{\text{Val}}[\mathcal{M}] + \varepsilon)}{\delta \min R}$  steps. Plugging this into (49) yields

$$E_1 \leq \frac{2\varepsilon(1 + \varepsilon/8 + \overline{\text{Val}}[\mathcal{M}])^2}{2(1 + \overline{\text{Val}}[\mathcal{M}])^2} \leq \varepsilon/4.$$

It follows that  $\text{Val}[\mathcal{M}^\star] - \text{Val}[\mathcal{M}, D(\delta, \bar{d})] \leq \frac{\varepsilon}{8} + \frac{\varepsilon}{4} \leq \frac{\varepsilon}{2}$ . It now suffices to show that  $\text{Val}[\mathcal{M}, D(\delta, \bar{d})] - \text{Val}[\mathcal{M}^\star] \leq \frac{\varepsilon}{2}$ .

To this end, let  $\mathbf{d}_\kappa$  be any  $(\varepsilon/8)$ -optimal strategy in  $\mathcal{M}_\kappa$ . Then  $\text{Bound}[\epsilon_\kappa, \mathbf{d}_\kappa] \leq \overline{\text{Val}}[\mathcal{M}_\kappa] + \varepsilon/8$ . Moreover, since  $\mathcal{M}_\kappa \in D(\delta, \bar{d})$ , using the same arguments as above (and the fact that rounding could only increase the one-step cost incurred) we get the same lower bound on one-step cost incurred by  $\mathbf{d}_\kappa$  and thus also the fact that  $\text{Bound}[\#, \mathbf{d}_\kappa] \leq \frac{2(\overline{\text{Val}}[\mathcal{M}] + \varepsilon)}{\delta \min R}$ . Plugging this into (50) and using the same computation as above, we get the desired inequality.  $\square$

## D Proofs for Section 4

**Proposition 5.** There is a number  $B \in \exp((\|C\| \cdot \|\underline{d}\| \cdot \bar{d})^{O(1)})$  such that for  $\delta = \varepsilon/B$  and  $\kappa = (\varepsilon \cdot \delta)/B$  it holds  $|\text{Val}[\mathcal{M}, D] - \text{Val}[\mathcal{M}_D^\star]| < \varepsilon$ .

Similarly to Section 3 we first bound the error caused by discretization and then by rounding:

**Lemma 17.** *Let*

$$\alpha := \frac{\varepsilon}{8 \cdot \text{Bound}[\epsilon, \mathbf{d}] \cdot \text{Bound}[\#, \mathbf{d}] \cdot |S'|}, \quad \delta := \alpha/D_1,$$

where  $D_1$  and is from Lemma 3. Then

$$\left| \text{Val}[\mathcal{M}] - \text{Val}[\mathcal{M}, D(\delta, \underline{d}, \bar{d}, \equiv)] \right| \leq \frac{\varepsilon}{2}.$$

**Lemma 18.** *Let  $\varepsilon > 0$  and fix  $\kappa = (\varepsilon/8 \cdot \text{Bound}[\epsilon, \mathbf{d}] \cdot \text{Bound}[\#, \mathbf{d}] \cdot |S'|)^2$ . Then it holds that*

$$\left| \text{Val}[\mathcal{M}, D(\delta, \underline{d}, \bar{d}, \equiv)] - \text{Val}[\mathcal{M}^\star] \right| \leq \frac{\varepsilon}{2}.$$

Note that from Lemma 17 and Lemma 18 directly follows Proposition 5.

### D.1 Proof of Lemma 17

To prove Lemma 17 we similarly to proof of Proposition 3 we will need bounds on expected number of steps and expected cost until reaching target. We will prove the following lemmas.

**Lemma 19.** *Let  $\mathbf{d}$  be a delay function such that for each  $s \in S'$  it holds that  $\underline{d} \leq \mathbf{d}(s) \leq \bar{d}$ . It holds that  $\text{Bound}[\epsilon, \mathbf{d}] \leq \left(\frac{|S_{\text{fd}}|}{\min P \cdot \min\{1, \lambda \cdot \underline{d}\}}\right)^{|S'|} \cdot e^{\lambda \bar{d} \cdot |S' \setminus S_{\text{fd}}|} \cdot (\max\{1/\lambda, \bar{d} \cdot \lambda + \bar{d}\} + 1) \cdot |S'| \cdot \max R$ .*

We first prove that there is a lower bound to reach a target from any state.

**Lemma 20.** *Let  $\mathcal{M}$  be DTMDP and  $\underline{d}, \bar{d} > 0$  be lower, upper bound on delay function, respectively. Let  $\mathbf{d}$  be any delay function such that for each state  $s \in S'$  it holds that  $\underline{d} \leq \mathbf{d}(s) \leq \bar{d}$ . Then for all  $s \in S'$  it holds that*

$$\Pr_{\mathcal{M}(\mathbf{d})}(\text{reach } G \text{ from } s \text{ in } |S' \setminus S_{\text{fd}}| \text{ steps}) \geq \left( \left( \frac{\min P}{|S_{\text{fd}}|} \right)^{|S_{\text{fd}}|} \cdot e^{-\lambda \bar{d}} \cdot \min\{1, (\lambda \cdot \underline{d})^{|S_{\text{fd}}|}\} \right)^{|S' \setminus S_{\text{fd}}|}$$

*Proof.* We first bound the minimal branching probability of  $T(s, |S_{\text{fd}}|/\lambda)$ . We use again uniformization method to compute the transient analysis of  $C'[s](\mathbf{d}(s))$  to evaluate  $T(s, |S_{\text{fd}}|/\lambda)$ . For all  $s \in S' \cap S_{\text{fd}}$  and  $s' \in S'$  that is reachable from  $\bar{s}$  in  $C'[s](\mathbf{d}(s))$  ( $\bar{s}$  is artificial initial state in  $C'[s](\mathbf{d}(s))$  corresponding to state  $s$ ) it holds that

$$\begin{aligned} T(s, \mathbf{d}(s))(s') &= e^{-\lambda \mathbf{d}(s)} \sum_{i=0}^{\infty} P^i(\bar{s}, s') \cdot \frac{(\lambda \cdot \mathbf{d}(s))^i}{i!} \\ &\geq e^{-\lambda \mathbf{d}(s)} \sum_{i=0}^{|S_{\text{fd}}|} P^i(\bar{s}, s') \cdot \frac{(\lambda \cdot \mathbf{d}(s))^i}{i!} \\ &\geq e^{-\lambda \bar{d}} \sum_{i=0}^{|S_{\text{fd}}|} P^i(\bar{s}, s') \cdot \frac{(\lambda \cdot \mathbf{d}(s))^i}{i!} \\ &\geq e^{-\lambda \bar{d}} \min_{i \in \{0, \dots, |S_{\text{fd}}|\}} \min P^{|S_{\text{fd}}|} \cdot \frac{(\lambda \cdot \mathbf{d}(s))^i}{i!} \\ &\geq e^{-\lambda \bar{d}} \min_{i \in \{0, \dots, |S_{\text{fd}}|\}} \min P^{|S_{\text{fd}}|} \cdot \frac{(\lambda \cdot \mathbf{d}(s))^i}{|S_{\text{fd}}|^{|S_{\text{fd}}|}} \\ &= \left( \frac{\min P}{|S_{\text{fd}}|} \right)^{|S_{\text{fd}}|} \cdot e^{-\lambda \bar{d}} \cdot \min_{i \in [0, |S_{\text{fd}}|]} (\lambda \cdot \underline{d})^i \\ &= \left( \frac{\min P}{|S_{\text{fd}}|} \right)^{|S_{\text{fd}}|} \cdot e^{-\lambda \bar{d}} \cdot \min\{1, (\lambda \cdot \underline{d})^{|S_{\text{fd}}|}\}, \end{aligned} \tag{51}$$

where (51) holds because the minimal positive branching probability in  $C'[s](\mathbf{d}(s))$  is  $\min P$  and each state, if it is reachable in  $C'[s](\mathbf{d}(s))$  from  $\bar{s}$ , it has to be reached by path of length at most  $|S_{\text{fd}}|$ . Thus the minimum discrete probability to reach any reachable state is at least  $\min P^{|S_{\text{fd}}|}$ . This probability is at least once multiplied by a factor  $\frac{(\lambda \cdot \mathbf{d}(s))^i}{i!}$  for appropriate  $i$ , thus we pick the smallest factor to get the correct lower bound on probability of moving from  $S' \cap S_{\text{fd}}$ . The minimal positive probability of moving from  $S' \setminus S_{\text{fd}}$  is simply  $\min P$

from definition of  $T(s, |S_{\text{fd}}|/\lambda)$ . Since  $\left(\frac{\min P}{|S_{\text{fd}}|}\right)^{|S_{\text{fd}}|} \cdot e^{-\lambda \bar{d}} \cdot \min\{1, (\lambda \cdot \underline{d})^{|S_{\text{fd}}|}\} \leq \min P$  it holds that for all  $s, s' \in S'$  such that  $T(s, |S_{\text{fd}}|/\lambda)(s') > 0$  implies

$$T(s, |S_{\text{fd}}|/\lambda)(s') \geq \left(\frac{\min P}{|S_{\text{fd}}|}\right)^{|S_{\text{fd}}|} \cdot e^{-\lambda \bar{d}} \cdot \min\{1, (\lambda \cdot \underline{d})^{|S_{\text{fd}}|}\}.$$

Now we provide a lower bound on probability of reaching  $G$  from any  $s \in S'$ :

$$\begin{aligned} \Pr_{\mathcal{M}(\mathbf{d})}(\text{reach } G \text{ from } s \text{ in } |S \setminus S_{\text{fd}}| \text{ steps}) &\geq (\min \{T(s', |S_{\text{fd}}|/\lambda)(s'') > 0 \mid s', s'' \in S'\})^{|S \setminus S_{\text{fd}}|} \\ &\geq \left(\left(\frac{\min P}{|S_{\text{fd}}|}\right)^{|S_{\text{fd}}|} \cdot e^{-\lambda \bar{d}} \cdot \min\{1, (\lambda \cdot \underline{d})^{|S_{\text{fd}}|}\}\right)^{|S \setminus S_{\text{fd}}|}. \end{aligned}$$

Now we are ready to prove Lemma 5.

*Proof.* We choose an arbitrary delay function  $\mathbf{d}$ , such that for all  $s \in S'$  holds  $\mathbf{d}(s) = d$  for some  $\underline{d} \leq d \leq \bar{d}$ . We compute the overestimation of expected cost until reaching goal states from any state  $s \in S'$ . During the computation we substitute  $\underline{d}$  or  $\bar{d}$  for  $d$  to get the correct upper bound.

We first bound the  $\mathbb{E}(s, \mathbf{d}(s))$  for each state  $s \in S'$ . Obviously for  $s \in S' \setminus S_{\text{fd}}$  the expected impulse cost paid in one step is  $\max R$  and the expected rate cost paid in one step is  $\max R/\lambda$ , thus together  $\mathbb{E}(s, \mathbf{d}(s)) \leq (1 + 1/\lambda) \cdot \max R$ . For state  $s \in S' \cap S_{\text{fd}}$  the expected number of exponential steps until reaching  $S'$  in time  $\mathbf{d}(s) = d$  is  $\lambda \cdot d \leq \lambda \cdot \bar{d}$  (employing mean of Poisson distribution). It is possible that one extra step is taken due to fixed-delay transition, thus the expected impulse cost is bounded by  $(\lambda \cdot \bar{d} + 1) \cdot \max R$ . Finally the expected rate cost paid until state from  $S'$  is reached from any  $S' \cap S_{\text{fd}}$  can be bounded from above by  $\mathbf{d}(s) \cdot \max R \leq \bar{d} \cdot \max R$ , what comprises of the maximum time  $\bar{d}$  until  $S'$  are reached (from definition of  $\mathbb{E}$  after firing of fixed-delay transition a state from  $S'$  is reached), and maximum rate cost bounded by  $\max R$ . Altogether we have:

$$\begin{aligned} \mathbb{E}(s, \mathbf{d}(s)) &\leq \max\{(1 + 1/\lambda) \cdot \max R, (\bar{d} \cdot \lambda + 1) \cdot \max R + \bar{d} \cdot \max R\} \\ &\leq (\max\{1/\lambda, \bar{d} \cdot \lambda + \bar{d}\} + 1) \cdot \max R. \end{aligned}$$

Now we bound the expected cost using Bernoulli trials, i.e. for all  $s \in S'$  it holds that

$$\begin{aligned} E_{\mathcal{M}[s](\mathbf{d})}^{\mathbb{E}} &\leq \sum_{i=1}^{\infty} i \cdot |S'| \cdot \max_{s'' \in S'} \mathbb{E}(s'', \mathbf{d}(s'')) \cdot (1 - pr)^{i-1} \cdot pr \\ &= \frac{|S'| \cdot \max_{s'' \in S'} \mathbb{E}(s'', \mathbf{d}(s''))}{pr} \\ &\leq \frac{|S'| \cdot (\max\{1/\lambda, \bar{d} \cdot \lambda + \bar{d}\} + 1) \cdot \max R}{\left(\left(\frac{\min P}{|S_{\text{fd}}|}\right)^{|S_{\text{fd}}|} \cdot e^{-\lambda \bar{d}} \cdot \min\{1, (\lambda \cdot \underline{d})^{|S_{\text{fd}}|}\}\right)^{|S \setminus S_{\text{fd}}|}} \\ &= \left(\frac{|S_{\text{fd}}|}{\min P \cdot \min\{1, \lambda \cdot \underline{d}\}}\right)^{|S|} \cdot e^{\lambda \bar{d} \cdot |S \setminus S_{\text{fd}}|} \cdot (\max\{1/\lambda, \bar{d} \cdot \lambda + \bar{d}\} + 1) \cdot |S'| \cdot \max R, \end{aligned}$$

where using Lemma 20

$$\begin{aligned} pr &= \min_{s'' \in S'} \Pr_{\mathcal{M}(\mathbf{d})}(\text{reach } G \text{ from } s'' \text{ in } |S \setminus S_{\text{fd}}| \text{ steps}) \\ &\geq \left( \left( \frac{\min P}{|S_{\text{fd}}|} \right)^{|S_{\text{fd}}|} \cdot e^{-\lambda \bar{d}} \cdot \min\{1, (\lambda \cdot \underline{d})^{|S_{\text{fd}}|}\} \right)^{|S \setminus S_{\text{fd}}|}. \end{aligned}$$

**Lemma 21.** Let  $\mathbf{d}$  be a delay function such that for each  $s \in S'$  it holds that  $\underline{d} \leq \mathbf{d}(s) \leq \bar{d}$ .

It holds that  $\text{Bound}[\#, \mathbf{d}] \leq \left( \frac{|S_{\text{fd}}|}{\min P \cdot \min\{1, \lambda \cdot \underline{d}\}} \right)^{|S|} \cdot e^{\lambda \bar{d} \cdot |S \setminus S_{\text{fd}}|} \cdot |S'|$ .

*Proof.* We use Bernoulli trials and Lemma 20 to bound the expected number of steps from any state to a goal state, i.e. for all  $s \in S'$  it holds that

$$\begin{aligned} E_{\mathcal{M}[\cdot](\mathbf{d})}^{\#} &\leq \\ &\leq \sum_{i=1}^{\infty} i \cdot |S'| \cdot (1 - \min_{s'' \in S'} \Pr_{\mathcal{M}(\mathbf{d})}(\text{reach } G \text{ from } s'' \text{ in } |S \setminus S_{\text{fd}}| \text{ steps}))^{i-1} \\ &\quad \cdot \min_{s'' \in S'} \Pr_{\mathcal{M}(\mathbf{d})}(\text{reach } G \text{ from } s'' \text{ in } |S \setminus S_{\text{fd}}| \text{ steps}) \\ &= \frac{|S'|}{\min_{s'' \in S'} \Pr_{\mathcal{M}(\mathbf{d})}(\text{reach } G \text{ from } s'' \text{ in } |S \setminus S_{\text{fd}}| \text{ steps})} \\ &\leq \frac{|S'|}{\left( \left( \frac{\min P}{|S_{\text{fd}}|} \right)^{|S_{\text{fd}}|} \cdot e^{-\lambda \bar{d}} \cdot \min\{1, (\lambda \cdot \underline{d})^{|S_{\text{fd}}|}\} \right)^{|S \setminus S_{\text{fd}}|}} \\ &= \left( \frac{|S_{\text{fd}}|}{\min P \cdot \min\{1, \lambda \cdot \underline{d}\}} \right)^{|S|} \cdot e^{\lambda \bar{d} \cdot |S \setminus S_{\text{fd}}|} \cdot |S'|. \end{aligned}$$

Now we have everything needed to prove Lemma 17

**Lemma 17.** Let

$$\alpha := \frac{\varepsilon}{8 \cdot \text{Bound}[\epsilon, \mathbf{d}] \cdot \text{Bound}[\#, \mathbf{d}] \cdot |S'|}, \quad \delta := \alpha / D_1,$$

where  $D_1$  is from Lemma 3. Then

$$\left| \text{Val}[\mathcal{M}] - \text{Val}[\mathcal{M}, D(\delta, \bar{d})] \right| \leq \frac{\varepsilon}{2}.$$

*Proof.* Please observe that Lemma 2 and Lemma 3 from Section 3 hold also for bounded optimization under partial observation. Moreover Lemma 3 is more general than we need here as we can use upper given bound  $\bar{d}$ . Now we just insert equations for  $\alpha$ ,  $\text{Bound}[\epsilon, \mathbf{d}]$ , and  $\text{Bound}[\#, \mathbf{d}]$  from Lemma 19 and Lemma 21 to the the statement of Lemma 2:

$$\begin{aligned} &2 \cdot \alpha \cdot \text{Bound}[\#, \mathbf{d}] \cdot (1 + \text{Bound}[\epsilon, \mathbf{d}] \cdot |S'|) \leq \\ &\leq 2 \cdot \frac{\varepsilon}{8 \cdot \text{Bound}[\epsilon, \mathbf{d}] \cdot \text{Bound}[\#, \mathbf{d}] \cdot |S'|} \cdot \text{Bound}[\#, \mathbf{d}] \cdot (1 + \text{Bound}[\epsilon, \mathbf{d}] \cdot |S'|) \\ &\leq \frac{\varepsilon}{4 \cdot \text{Bound}[\epsilon, \mathbf{d}] \cdot |S'|} \cdot 2 \cdot \text{Bound}[\epsilon, \mathbf{d}] \cdot |S'| \\ &\leq \frac{\varepsilon}{2}. \end{aligned}$$

Note that for  $B$  from Proposition 5 holds

$$\begin{aligned} B &= 8 \cdot \text{Bound}[\epsilon, \mathbf{d}] \cdot \text{Bound}[\#, \mathbf{d}] \cdot |S'| \\ &= 8 \cdot \left( \left( \frac{|S_{\text{fd}}|}{\min P \cdot \min\{1, \lambda \cdot d\}} \right)^{|S|} \cdot e^{\lambda \bar{d} \cdot |S \setminus S_{\text{fd}}|} \right)^2 \cdot |S'|^3 \cdot (\max\{1/\lambda, \bar{d} \cdot \lambda + \bar{d}\} + 1) \cdot \max R. \end{aligned}$$

Obviously  $B \in \exp((\|C\| \cdot \|\underline{d}\| \cdot \bar{d})^{O(1)})$ .

## D.2 Proof of Lemma 18

**Lemma 18.** Let  $\varepsilon > 0$  and fix  $\kappa = (\varepsilon/8 \cdot \text{Bound}[\epsilon, \mathbf{d}] \cdot \text{Bound}[\#, \mathbf{d}] \cdot |S'|)^2$ . Then it holds that

$$\left| \text{Val}[\mathcal{M}, D(\delta, \underline{d}, \bar{d}, \equiv)] - \text{Val}[\mathcal{M}^*] \right| \leq \frac{\varepsilon}{2}.$$

*Proof.* We proceed similarly to proof of Proposition 4. The only difference is that from Lemma 19 and Lemma 21 we have bounds for arbitrary delay functions  $\mathbf{d}, \mathbf{d}_\kappa \in D(\underline{d}, \bar{d}, \equiv)$ , thus the argumentation is much simpler. We show only equation (49) (for equation (50) the argumentation is symmetric):

$$\begin{aligned} E_1 &= 2 \cdot \kappa \cdot \text{Bound}[\#, \mathbf{d}] (1 + \text{Bound}[\epsilon, \mathbf{d}] \cdot |S'|) \\ &\leq 2 \cdot \left( \frac{\varepsilon}{8 \cdot \text{Bound}[\epsilon, \mathbf{d}] \cdot \text{Bound}[\#, \mathbf{d}] \cdot |S'|} \right)^2 \cdot \text{Bound}[\#, \mathbf{d}] (1 + \text{Bound}[\epsilon, \mathbf{d}] \cdot |S'|) \\ &\leq 2 \cdot \frac{\varepsilon}{16 \cdot \text{Bound}[\epsilon, \mathbf{d}] \cdot \text{Bound}[\#, \mathbf{d}] \cdot |S'|} \cdot \text{Bound}[\#, \mathbf{d}] \cdot (1 + \text{Bound}[\epsilon, \mathbf{d}] \cdot |S'|) \\ &\leq \frac{\varepsilon}{8 \cdot \text{Bound}[\epsilon, \mathbf{d}] \cdot |S'|} \cdot 2 \cdot \text{Bound}[\epsilon, \mathbf{d}] \cdot |S'| \\ &\leq \frac{\varepsilon}{4}. \end{aligned}$$

Note that for  $B$  from Proposition 5 holds

$$\begin{aligned} B &= 8 \cdot \text{Bound}[\epsilon, \mathbf{d}] \cdot \text{Bound}[\#, \mathbf{d}] \cdot |S'| \\ &= 8 \cdot \left( \left( \frac{|S_{\text{fd}}|}{\min P \cdot \min\{1, \lambda \cdot d\}} \right)^{|S|} \cdot e^{\lambda \bar{d} \cdot |S \setminus S_{\text{fd}}|} \right)^2 \cdot |S'|^3 \cdot (\max\{1/\lambda, \bar{d} \cdot \lambda + \bar{d}\} + 1) \cdot \max R. \end{aligned}$$

Obviously  $B \in \exp((\|C\| \cdot \|\underline{d}\| \cdot \bar{d})^{O(1)})$ .

## D.3 Proof of Correctness of the NP-Algorithm in Unary Case

Let us show that the presented algorithm is correct for the problem in Theorem 3. First, assume that  $\text{Val}[C, D] < x - \varepsilon$ . The optimal (MD) strategy  $\mathbf{d}$  in  $\mathcal{M}_D^*$  from the set  $D(\underline{d}, \bar{d}, \equiv)$  hence satisfies  $E_{\mathcal{M}_D^*(\mathbf{d})} < x$ , due to Proposition 5. The algorithm accepts as it non-deterministically guesses  $\mathbf{d}$ . Similarly if  $\text{Val}[C, D] > x + \varepsilon$ , no strategy from the set  $D$  guarantees in  $\mathcal{M}_D^*$  expected total cost  $\leq x$ , hence the algorithm does not accept. This concludes the proof that the problem from Theorem 3 is in NP.

#### D.4 Proof of Proposition 6 (NP Hardness)

**Proposition 6.** For a formula  $\varphi$  in CNF with  $k$  literals,  $C_\varphi$  and  $Cost_\varphi$  are constructed in time polynomial in  $k$  and, furthermore,

$$Val[C_\varphi, D] < 17k^2 \text{ if } \varphi \text{ is satisfiable} \quad \text{and} \quad Val[C_\varphi, D] > 17k^2 + 1, \text{ otherwise.}$$

*Proof.* If  $\varphi$  is satisfiable, let  $\nu$  denote the satisfying truth assignment. Recall that we set  $\underline{d} = 0.01$  and  $\bar{d} = 16k$ . We set  $\mathbf{d}(s_{in}) := \underline{d}$  and  $\mathbf{d}(s_{i,j}^0) := \bar{d}$  if  $\nu(X) = 1$  and  $\mathbf{d}(s_{i,j}^0) := \underline{d}$  if  $\nu(X) = 0$ , where  $X$  is the variable of the literal  $l_{i,j}$  and arbitrarily in other states. In  $C_\varphi(\mathbf{d})$  in the component for any TRUE literal  $l_{i,j}$ , i.e. with  $\nu(l_{i,j}) = 1$ , the goal is reached from  $s_{i,j}^0$  with probability  $> 0.99$  before leaving the component. Indeed, the probability to take no exponential transition within time  $0.01$  is  $> 0.99$  and the probability to take at least  $8k$  exponential transitions within time  $16k$  is  $> 0.99$  for any  $k \in \mathbb{N}$ . As each clause  $\varphi_i$  has at most  $k$  literals and at least one TRUE literal, the expected cost incurred in the component for  $\varphi_i$  is at most  $(16k \cdot k)/0.99 < 17k^2$ .

As regards the other implication, let  $\varphi$  be not satisfiable such that  $k \geq 7$  (note that we can assume it as when fixing the number of literals, the SAT problem becomes polynomial) and let  $\mathbf{d} \in D(\underline{d}, \bar{d}, \equiv)$ . We need to show that  $E_{C_\varphi(\mathbf{d})} > 17k^2 + 1$ . Based on  $\mathbf{d}$ , we construct a truth assignment  $\nu$  such that  $\nu(X) = 0$  iff for all literals  $l_{i,j}$  with variable  $X$ , we have  $\mathbf{d}(s_{i,j}) \leq k$ . Since  $\varphi$  is unsatisfiable, there must be at least one clause  $i$  that is not satisfied w.r.t.  $\nu$ . It suffices to show that for the initial state of the component of the  $i$ -th clause we have  $E_i := E_{C_\varphi[s_{i,1}^0]}(\mathbf{d}) > k \cdot (17k^2 + 1)$ . Indeed, even for the unrealistic worst case that the expected cost in components of other clauses is 0, the overall expected cost is still  $> 17k^2 + 1$ .

We know that all negative literals have the delay above  $k \geq 7$  and all the positive literals have the delay below  $k$ . Let us assume the worst case that all the delays are  $k$ . With delay equal  $k$ , let us by  $p_{pos}$  and  $p_{neg}$  denote the probabilities that from the initial state of a component for a positive and negative literal, respectively, the goal is reached before leaving the component. It is easy to show that for all  $k \geq 7$ :

$$p_{pos} := 1 - e^{-k} \cdot \sum_{n=0}^{8k-1} \frac{k^n}{n!} \leq \frac{1}{17k^2 + 1}, \text{ and} \quad p_{neg} := e^{-k} \leq \frac{1}{17k^2 + 1}.$$

Based on this observation, we can under-approximate  $E_i$  by

$$E_i \geq \frac{k}{\frac{1}{17k^2 + 1}} = k \cdot (17k^2 + 1).$$

□